

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

# OM protein - protein search, using SW model

Run on: October 22, 2003, 12:19:30 : Search time 86 Seconds

(without alignments)  
699,504 Million cell updates/sec

Title: US-09-805-311-2

Perfect score: 1939

Sequence: 1 MGICG1TKLADNAPKAKE.....SDKTSKAAANKKTKAGKKK 379

Scoring table:

BLOSUM62  
Gapop 10.0 , Capext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08  
Maximum Match 1008

Listing first 45 summaries

Database :

A\_Geneseq\_19Jun03:\*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1939	100.0	379	21	AAV95307
2	1939	100.0	379	21	AAV95309
3	1939	99.7	379	21	AAV95308
4	1939	99.7	379	21	AAV95310
5	1031	53.2	380	20	AAW92504
6	1007.5	52.0	377	20	AAW92505
7	998.5	51.5	385	22	ABB63960
8	996	51.4	378	20	AAW92508
9	938	48.4	382	20	AAW92505

10	699.5	36.1	373	22	ABG19545	Novel human di...
11	648.5	33.4	340	18	AAW4216	Pyrococcus fur...
12	648.5	33.4	340	19	AAW9970	Pyrococcus fur...
13	648.5	33.4	340	19	AAW59440	Amino acid seq...
14	648.5	33.4	340	23	ABG52941	Pyrococcus fur...
15	647	33.4	326	19	AAW59953	Amino acid seq...
16	644	33.2	332	19	AAW59951	Amino acid seq...
17	636	32.8	343	22	AAW59944	Putative P. aby...
18	632.5	32.6	343	20	AAW59778	Pyrococcus heat...
19	609	31.4	325	19	AAW59950	Amino acid seq...
20	600	30.9	340	19	AAW59949	Amino acid seq...
21	562	29.0	340	19	AAW59952	Amino acid seq...
22	560.5	28.9	326	18	AAW54215	Methanococcus j...
23	560.5	28.9	326	19	AAW59939	Amino acid seq...
24	560.5	28.9	326	23	ABG59942	Methanococcus j...
25	553.5	28.5	326	19	AAW79969	Methanococcus j...
26	538.5	27.8	336	19	AAW79982	Archaeoglobus v...
27	538.5	27.8	336	19	AAW59946	Amino acid seq...
28	538.5	27.8	336	23	ABG59944	Archaeoglobus v...
29	532	27.4	338	23	ABG59945	Archaeoglobus v...
30	525	27.1	328	19	AAW59948	Archaeoglobus v...
31	525	27.1	328	23	ABG59947	Archaeoglobus v...
32	411	21.2	258	19	AAW59947	Methanobacteri...
33	340.5	17.6	386	20	AAW42507	Amino acid seq...
34	336	17.3	1516	21	AAW18195	Yeast delta-RAU...
35	311	16.0	96	23	ABP3304	Plasmidium fal...
36	257	13.3	726	22	ABW5192	Human nuclease
37	241.5	12.6	1236	22	ABW5297	Drosophila mel...
38	241.5	12.6	1257	22	ABW6127	Drosophila mel...
39	215.5	11.1	871	23	AAW4927	Thermophilic D...
40	215.5	11.1	871	23	AAW4928	Thermophilic D...
41	210	10.8	732	22	ABW7117	Drosophila mel...
42	208	10.7	872	20	AAW31815	Thermococcus ba...
43	205	10.6	872	18	AAW56805	Thermococcus ba...
44	205	10.6	872	20	AAW51812	Thermococcus ba...
45	205	10.6	872	20	AAW31816	Thermohydro...

## ALIGNMENTS

RESULT 1	AAV95307	standard: Protein: 379 AA.
ID	AAV95307	
XX	AAV95307	
AC	AAV95307	
XX		
DT	12-SEP-2000	(first entry)
XX		
DE	Maize Rad2/FEN-1 protein.	
XX		
KW	Maize; Rad2/FEN-1; transgenic plant; male sterile plant; endonuclease; exonuclease; DNA repair; gene targeting.	
XX		
OS	Zea mays.	
XX		
PN	W0200036109-A1.	
XX		
PC	22-JUN-2000.	
XX		
PF	16-NOV-1999;	99WO-US27147.
XX		
FR	15-DEC-1998;	98US-0112332.
XX		
FA	(PION-) PIONEER HI-BRED INT INC.	
PI	Mahajan Pa.	
LR	WPI: 2000-452026/39.	
DR	N-PSDB: AAA27923.	
XX		
PT	Maize RAD2/FEN-1 nucleic acids and proteins useful for modulating DNA recombination and repair in transgenic plants, e.g. for gene targeting	

PT and the production of male sterile plants -  
 PS Claim 11: Page 71-72: 85pp: English.

CC The present sequence is that of maize RAD2/FEN-1, as deduced from  
 CC a cDNA clone (see AAA27923) isolated from maize line B73 immature  
 CC ear tissue. Rad2/FEN-1 is a structure specific endonuclease which  
 CC under certain conditions also acts as an exonuclease. Rad2/FEN-1  
 CC can be expressed in transgenic plant cells using conventional  
 CC methods. The protein is involved in the regulation of DNA repair  
 CC and recombination in plant systems and therefore may be used for  
 CC improving gene targeting during further recombinant DNA protocols  
 CC involving plants. RAD2/FEN-1 endonucleolytic activity is essential  
 CC in DNA replication and nucleotide excision and repair reactions.  
 CC The exolytic activity is involved in double strand break repair and  
 CC end joining. The protein is also useful in strand exchange  
 CC reactions during homologous recombination. These functions may be  
 CC useful in gene targeting and in the production of male sterile  
 CC plants. The efficacy of gene targeting can be improved by the  
 CC overexpression of exogenous Rad2/FEN-1 while male sterile plants can  
 CC be produced by the down regulation of Rad2/FEN-1 expression.

CC Sequence 379 AA:

Query Match 100.0%; Score 1939; DB 21; Length 379;

Best Local Similarity 100.0%; Pred. No. 1,3e-170; Mismatches 0; Indels 0; Gaps 0;

Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGKGLTKLLADNAPKAKKEQKESYFGKRIADVASKSIYOFLLVGRGTMETLTNEAGE 60

DB 1 MGKGLTKLLADNAPKAKKEQKESYFGKRIADVASKSIYOFLLVGRGTMETLTNEAGE 60

QY 61 VTSHLQGMFNRTIRLLEAGIKPVYVFDGKPPDMKKOELAKRYSKRDATKDLTEAVEVGD 120

DB 61 VTSHLQGMFNRTIRLLEAGIKPVYVFDGKPPDMKKOELAKRYSKRDATKDLTEAVEVGD 120

QY 121 KDAIEKLSKRTYKVRQHNEDCKRLRLMGVYVVEASEAECAACINDKVFVAASED 180

DB 121 KDAIEKLSKRTYKVRQHNEDCKRLRLMGVYVVEASEAECAACINDKVFVAASED 180

QY 181 MDSLTFGAPRFLRHLMPPSSKKIPVMEFDVAKVLELELTMDQFDLCTLCGGDYCDSIK 240

DB 181 MDSLTFGAPRFLRHLMPPSSKKIPVMEFDVAKVLELELTMDQFDLCTLCGGDYCDSIK 240

QY 241 GIGGOTALKLIRHGSLESTLENLNKDRYOQPEDMVQOARLKEKRNVTLDPELKMTA 300

DB 241 GIGGOTALKLIRHGSLESTLENLNKDRYOQPEDMVQOARLKEKRNVTLDPELKMTA 300

QY 301 PDEGLISFLVKNQNGNEDRYTKAIEKISAKKSSQGRLESEFKPTATTSAPLKRKETS 360

DB 301 PDEGLISFLVKNQNGNEDRYTKAIEKISAKKSSQGRLESEFKPTATTSAPLKRKETS 360

QY 361 DKTSAKAANKKTKAGGKKK 379

DB 361 DKTSAKAANKKTKAGGKKK 379

RESULT 2

ID AAY95309 standard; Protein: 379 AA.

AC AAY95309:

UT 12-SEP-2000 (first entry)

DE Maize Rad2/FEN-1 protein.

KW Maize: Rad2/FEN-1: transgenic plant; male sterile plant;

KM exonuclease; exonuclease; DNA repair; gene targeting.

OS Zea mays.

PN W020C036109-A1.

XX 22-JUN-2000.

XX 16-NOV-1999; 39MO-US27147.

XX 15-DEC-1998; 390US-0112332.

XX (PION-) PIONEER HI-BRED INT INC.

XX Mahajan PH:

XX WPI: 2000-452026/39.

XX N-PSDB; AAA27925.

XX Maize RAD2/FEN-1 nucleic acids and proteins useful for modulating DNA

XX recombination and repair in transgenic plants, e.g. for gene targeting

XX and the production of male sterile plants -

XX Example 1: Page 76-77: 85pp: English.

XX The present sequence is that of maize RAD2/FEN-1, as deduced from

XX a cDNA clone (see AAA27925) derived from maize line W23 tassel

XX polyA RNA. Rad2/FEN-1 is a structure specific endonuclease which

XX under certain conditions also acts as an exonuclease. Rad2/FEN-1

XX can be expressed in transgenic plant cells using conventional

XX methods. The protein is involved in the regulation of DNA repair

XX and recombination in plant systems and therefore may be used for

XX improving gene targeting during further recombinant DNA protocols

XX involving plants. RAD2/FEN-1 endonucleolytic activity is essential

XX in DNA replication and nucleotide excision and repair reactions.

XX The exolytic activity is involved in double strand break repair and

XX end joining. The protein is also useful in strand exchange

XX reactions during homologous recombination. These functions may be

XX useful in gene targeting and in the production of male sterile

XX plants. The efficacy of gene targeting can be improved by the

XX overexpression of exogenous Rad2/FEN-1 while male sterile plants can

XX be produced by the down regulation of Rad2/FEN-1 expression.

XX Sequence 379 AA:

XX Query Match 100.0%; Score 1939; DB 21; Length 379;

XX Best Local Similarity 100.0%; Pred. No. 1,3e-170; Mismatches 0; Indels 0; Gaps

XX Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps

QY 1 MGKGLTKLLADNAPKAKKEQKESYFGKRIADVASKSIYOFLLVGRGTMETLTNEAGE 60

DB 1 MGKGLTKLLADNAPKAKKEQKESYFGKRIADVASKSIYOFLLVGRGTMETLTNEAGE 60

QY 61 VTSHLQGMFNRTIRLLEAGIKPVYVFDGKPPDMKKOELAKRYSKRDATKDLTEAVEVGD 120

DB 61 VTSHLQGMFNRTIRLLEAGIKPVYVFDGKPPDMKKOELAKRYSKRDATKDLTEAVEVGD 120

QY 121 KDAIEKLSKRTYKVRQHNEDCKRLRLMGVYVVEASEAECAACINDKVFVAASED 180

DB 121 KDAIEKLSKRTYKVRQHNEDCKRLRLMGVYVVEASEAECAACINDKVFVAASED 180

QY 181 MDSLTFGAPRFLRHLMPPSSKKIPVMEFDVAKVLELELTMDQFDLCTLCGGDYCDSIK 240

DB 181 MDSLTFGAPRFLRHLMPPSSKKIPVMEFDVAKVLELELTMDQFDLCTLCGGDYCDSIK 240

QY 241 GIGGOTALKLIRHGSLESTLENLNKDRYOQPEDMVQOARLKEKRNVTLDPELKMTA 300

DB 241 GIGGOTALKLIRHGSLESTLENLNKDRYOQPEDMVQOARLKEKRNVTLDPELKMTA 300

QY 301 PDEGLISFLVKNQNGNEDRYTKAIEKISAKKSSQGRLESEFKPTATTSAPLKRKETS 360

DB 301 PDEGLISFLVKNQNGNEDRYTKAIEKISAKKSSQGRLESEFKPTATTSAPLKRKETS 360

QY 361 DKTSAKAANKKTKAGGKKK 379

DB 361 DKTSAKAANKKTKAGGKKK 379



Best local similarity 99.7%, Pred. No. 4,7e-170;  
 Matches 378; Conservative 0; Mismatches 1; Indels 0; Gaps 0

QY 1 MGKGLTKLLADNAPKAKKEQKESFYGRKIVADSNISYQFLVGRISKFTLLNEAGS 60  
 |||||  
 DB 1 MGKGLTKLLADNAPKAKKEQKESFYGRKIVADSNISYQFLVGRISKFTLLNEAGS 60  
 |||||

QY 61 VTSHLQGMFNRTIRLLFAGIKPVYVFDGKPPDMKKQELARYSKRDATKLTAEVGD 120  
 |||||  
 DB 61 VTSHLQGMFNRTIRLLFAGIKPVYVFDGKPPDMKKQELARYSKRDATKLTAEVGD 120  
 |||||

QY 121 KDAIEKLSKRTYKVTROHNEDECKRLRLMGVYVFAESEAFAECALCINQYFAVASD 180  
 |||||  
 DB 121 KDAIEKLSKRTYKVTROHNEDECKRLRLMGVYVFAESEAFAECALCINQYFAVASD 180  
 |||||

QY 181 MDLTFGAPRFLRLHMDPSSKKIPVMEFDYAKVLEELTMDQFIDLCILGCDYCSIR 240  
 |||||  
 DB 181 MDLTFGAPRFLRLHMDPSSKKIPVMEFDYAKVLEELTMDQFIDLCILGCDYCSIR 240  
 |||||

QY 241 GIGGOTALKLIRHOGSTESILENKKDRYOIPEDMPYQAFARLFEKPNVTLDIPELKWT 300  
 |||||  
 DB 241 GIGGOTALKLIRHOGSTESILENKKDRYOIPEDMPYQAFARLFEKPNVTLDIPELKWT 300  
 |||||

QY 301 PDEEGLISFLVKGDFNEDRVTKAIEKIKSAKNSOGRLSEFFKPTATTSAPLKRKETS 360  
 |||||  
 DB 301 PDEEGLISFLVKGDFNEDRVTKAIEKIKSAKNSOGRLSEFFKPTATTSAPLKRKETS 360  
 |||||

QY 361 DKTSKAANKTKKAGKKK 379  
 |||||  
 DB 361 DKTSKAANKTKKAGKKK 379  
 |||||

RESULT 5  
 AAM92504  
 ID AAM92504 standard; Protein: 380 AA.

AC AAM92504:  
 DT 23-APR-1999 (first entry)  
 XX Human FEN-1 protein.  
 DE Human FEN-1 protein.  
 XX FEN-1; human; flap endonuclease; detection; diagnosis; carcinogen;  
 KM neoplasma; antineoplastic agent; cleavage.  
 XX Homo sapiens.  
 OS Homo sapiens.  
 PN J55874283-A.  
 XX 23-APR-1999.  
 PD 23-APR-1999.  
 XX 30-MAY-1995; 95US-0455968.  
 PF 30-MAY-1995; 95US-0455968.  
 PR 30-MAY-1995; 95US-0455968.  
 XX (HARR/) HARRINGTON J J.  
 PA (HSIE/) HSIEH C.  
 PA (LIEB/) LIEBER M R.  
 XX Harrington JJ, Hsieh C, Lieber MR;  
 PT WPI: 1999-179985/15.  
 DR N-PSDB: AAX02107.  
 XX DNA encoding flap endonuclease polypeptides - useful for producing  
 PT e.g. recombinant polypeptides  
 XX Claim 1; Fig 1A: 58pp; English.  
 PS This sequence represents a human FEN-1 (flap endonuclease) protein. This  
 CC protein can be used in methods for detecting a pathological condition in  
 CC a patient, for diagnostic purposes, for screening for antineoplastic  
 CC agents and carcinogens, for diagnostic staging of neoplasia, for

CC producing recombinant flap endonuclease for use as research or  
 CC diagnostic reagents, for producing antibodies reactive with the novel  
 CC polypeptides, for producing transgenic nonhuman animals expressing the  
 CC novel polypeptides encoded by a transgene. The invention also provides  
 CC novel molecular cloning techniques and reagents involving cleavage of  
 CC a flap or nick with a flap endonuclease.

50 Sequence 380 AA:  
 Query Match 53.2%; Score 1031; DB 20; Length 380;  
 Best local similarity 53.8%; Pred. No. 1.2e-86;  
 Matches 206; Conservative 69; Mismatches 98; Indels 10; Gaps

QY 1 MGKGLTKLLADNAPKAKKEQKESFYGRKIVADSNISYQFLVGRISKFTLLNEAGS 60  
 |||||  
 DB 1 MGKGLTKLLADNAPKAKKEQKESFYGRKIVADSNISYQFLVGRISKFTLLNEAGS 60  
 |||||

QY 61 VTSHLQGMFNRTIRLLFAGIKPVYVFDGKPPDMKKQELARYSKRDATKLTAEVGD 120  
 |||||  
 DB 61 VTSHLQGMFNRTIRLLFAGIKPVYVFDGKPPDMKKQELARYSKRDATKLTAEVGD 120  
 |||||

QY 121 KDAIEKLSKRTYKVTROHNEDECKRLRLMGVYVFAESEAFAECALCINQYFAVASD 180  
 |||||  
 DB 121 KDAIEKLSKRTYKVTROHNEDECKRLRLMGVYVFAESEAFAECALCINQYFAVASD 180  
 |||||

QY 181 MDLTFGAPRFLRLHMDPSSKKIPVMEFDYAKVLEELTMDQFIDLCILGCDYCSIR 240  
 |||||  
 DB 181 MDLTFGAPRFLRLHMDPSSKKIPVMEFDYAKVLEELTMDQFIDLCILGCDYCSIR 240  
 |||||

QY 241 GIGGOTALKLIRHOGSTESILENKKDRYOIPEDMPYQAFARLFEKPNVTLDIPELKWT 300  
 |||||  
 DB 241 GIGGOTALKLIRHOGSTESILENKKDRYOIPEDMPYQAFARLFEKPNVTLDIPELKWT 300  
 |||||

QY 301 PDEEGLISFLVKGDFNEDRVTKAIEKIKSAKNSOGRLSEFFKPTATTSAPLKRKETS 360  
 |||||  
 DB 301 PDEEGLISFLVKGDFNEDRVTKAIEKIKSAKNSOGRLSEFFKPTATTSAPLKRKETS 360  
 |||||

QY 361 DKTSKAANKTKKAGKKK 379  
 |||||  
 DB 361 DKTSKAANKTKKAGKKK 379  
 |||||

RESULT 6  
 AAM92505  
 ID AAM92505 standard; Protein: 377 AA.

AC AAM92505:  
 DT 23-APR-1999 (first entry)  
 XX Mouse FEN-1 protein.  
 DE Mouse FEN-1 protein.  
 XX FEN-1; mouse; flap endonuclease; detection; diagnosis; carcinogen;  
 KM neoplasma; antineoplastic agent; cleavage.  
 XX Mus sp.  
 OS Mus sp.  
 PN U05874283-A.  
 XX 23-APR-1999.  
 PD 23-APR-1999.  
 XX 30-MAY-1995; 95US-0455968.  
 PF 30-MAY-1995; 95US-0455968.  
 PR 30-MAY-1995; 95US-0455968.  
 XX (HARR/) HARRINGTON J J.  
 PA (HSIE/) HSIEH C.  
 PA (LIEB/) LIEBER M R.  
 XX Harrington JJ, Hsieh C, Lieber MR;  
 PT WPI: 1999-179985/15.  
 DR N-PSDB: AAX02108.  
 XX DNA encoding flap endonuclease polypeptides - useful for producing  
 PT e.g. recombinant polypeptides  
 XX Claim 1; Fig 1A: 58pp; English.  
 PS This sequence represents a human FEN-1 (flap endonuclease) protein. This  
 CC protein can be used in methods for detecting a pathological condition in  
 CC a patient, for diagnostic purposes, for screening for antineoplastic  
 CC agents and carcinogens, for diagnostic staging of neoplasia, for

XX DNA encoding flap endonuclease polypeptides - useful for producing  
 PT e.g. recombinant polypeptides  
 XX  
 XX  
 PS Claim 1: Fig 2A: 58pp; English.  
 CC This sequence represents a mouse FEN-1 (flap endonuclease) protein. This  
 CC protein can be used in methods for detecting a pathological condition in  
 CC a patient, for diagnostic purposes, for screening for antineoplastic  
 CC agents and carcinogens, for diagnostic staging of neoplasia, for  
 CC producing recombinant flap endonuclease for use as research or  
 CC diagnostic reagents, for producing antibodies reactive with the novel  
 CC polypeptides, for producing transgenic nonhuman animals expressing the  
 CC novel polypeptides encoded by a transgene. The invention also provides  
 CC novel molecular cloning techniques and reagents involving cleavage of  
 CC a flap or nick with a flap endonuclease.  
 XX  
 XX  
 SQ Sequence 377 AA:  
 Query Match 52.0%; Score 1007.5; DB 20; Length 377;  
 Best Local Similarity 53.4%; Pred. No. 1,86-84;  
 Matches 203; Conservative 68; Mismatches 100; Indels 9; Gaps 7;  
 QY 1 MGIGLTKLADNAPKAKKEQFESYFGKRIADVDSMSIYQFLIVGRTGMEITJNEAGE 60  
 DB 1 MGIGLAKLADVAPSAIRENDIKSYFGKRVADMSIYQFLIAY-ROGCDVLOQNEDE 59  
 QY 61 VSHOGMNFRTIRLEGIKPYVFDGKPPDMKKOELAKRYSKRDATKDLTEAVSGD 120  
 DB 60 TTS-LMGMEYRTIR-MENGIRPYVFDGKPPDLKSGELAKSERARAEKQADQADEAG- 116  
 QY 121 KDAIEKSKRTYKVTROHNEDECKRLRLMGVPPVAPSEAEAECAALCINDKPVAVASED 180  
 DB 117 MEEVEKFTKRLVKYTKQHNDECKHLLSLMGIPYLDAPSEASCAALAKGKYVAAATD 176  
 QY 181 MDLTFGAPRFLRHLDMPSSKKIPMEFDVAVYLELELTMDQITDLCILCCDYSISIK 240  
 DB 177 MDCLTFGSPVLMRLTLTASEAKLPIQEFHLSRVLOELTINQOEVDCILCLGSDYCESIR 236  
 QY 241 GIGGOTALKLIRHGSISLELNLNDRYOIPEDMPYOEARLFEKPNVT-LDIPELKMT 299  
 DB 237 GIGAKRVALLIOKHSIEIYVRIDPSKYPPENMLHKEAQDLFEPEVPESEVAKMS 296  
 QY 300 APDEGILISFLVKGDFNEDRYTKAIEKISAKKSSOGRLSEFPKPAITSAFLKHKET 359  
 DB 297 EPNEBEIVKFMCGFQSFSEIRISGVKRLSKSRGOSTGRLDDEFKYGSLSS-AKRKEP 355  
 QY 360 SDKTSKAANKKTKAGGKK 379  
 DB 356 EPKGS---AKKAKTGAGAK 372

RESULT 7  
 ABB63960  
 ID ABB63960 standard; Protein: 385 AA.  
 AC ABB63960:  
 XX  
 XX 26-MAR-2002 (first entry)  
 XX  
 XX Drosophila melanogaster polypeptide SEQ ID NO 18672.  
 DE  
 XX  
 XX Drosophila: developmental biology; cell signalling; insecticide;  
 KM pharmaceutical.  
 XX  
 XX Drosophila melanogaster.  
 OS  
 XX  
 XX MO200171042-A2.  
 PN  
 XX  
 XX 27-SEP-2001.  
 PD  
 XX  
 XX 23-MAR-2001; 2001WO-US09231.  
 PE  
 XX

PR 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 XX  
 XX (PEKE ) PE CORP NY.  
 PA  
 XX  
 XX Venter JC, Adams M, Li PWD, Myers EW;  
 PI  
 XX  
 XX WPI; 2001-656860/75.  
 DR  
 XX  
 XX N-PSDB; ABI08063.  
 CC  
 XX  
 XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 PS Disclosure; SEQ ID NO 18672; 21pp + Sequence listing; English.  
 CC  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABU1840-ABU3411), expressed DNA  
 CC sequences (ABU1840-ABU3411) and the encoded proteins  
 CC (ABU1840-ABU3411).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIP  
 CC at [http://wipo.int/pub/published\\_pct\\_sequences](http://wipo.int/pub/published_pct_sequences).  
 XX  
 XX  
 SQ Sequence 385 AA:  
 Query Match 51.5%; Score 999.5; DB 22; Length 385;  
 Best Local Similarity 52.9%; Pred. No. 1,36-83;  
 Matches 202; Conservative 65; Mismatches 104; Indels 11; Gaps  
 QY 1 MGIGLTKLADNAPKAKKEQFESYFGKRIADVDSMSIYQFLIVGRTGMEITJNEAGE  
 DB 1 MGIGLTKLADNAPKAKKEQFESYFGKRIADVDSMSIYQFLIVGRTGMEITJNEAGE  
 QY 61 VSHOGMNFRTIRLEGIKPYVFDGKPPDMKKOELAKRYSKRDATKDLTEAVSGD  
 DB 60 TTS-LMGMEYRTIR-LDNGIKPYVFDGKPPDLKSGELAKSERARAEKQADQADEAG-  
 QY 121 KDAIEKSKRTYKVTROHNEDECKRLRLMGVPPVAPSEAEAECAALCINDKPVAVASED  
 DB 120 DAGIEKFNRLVAVTKEHAKKELLTLGVPVADPCAEKQCAALVAGKYATATATED  
 QY 181 MDLTFGAPRFLRHLDMPSSKKIPMEFDVAVYLELELTMDQITDLCILCCDYSISIK  
 DB 180 MDALTFGSTKRLYLYSEARKKPYVEFYDKLTLEGLAINNREFIDLCILCCDYSISIK  
 QY 241 GIGGOTALKLIRHGSISLELNLNDRYOIPEDMPYOEARLFEKPNVT-LDIPELKMT  
 DB 240 GIGPKRAIETINTYROIETILLDLSSKYTVPENMVYKARELFEPEVADSDILKAV  
 QY 300 APDEGILISFLVKGDFNEDRYTKAIEKISAKKSSOGRLSEFPKPAITSAFLKHKET  
 DB 303 EPDEBGIVKFLCGDROFNEDRYVNGAKKILKSKGATQVRLDSFFKILPSTFNATIAA--  
 QY 355 KKEKTSKTSKAANKKTKAGG 375  
 DB 358 KKK--AKKAKKSNKKAKKTSG 377

RESULT 8  
 AAW92508  
 ID AAW92508 standard; Protein: 378 AA.  
 AC AAW92508:  
 XX  
 XX  
 XX 23-APR-1999 (first entry)  
 XX  
 XX Human FEN-1 protein.  
 DE  
 XX



QY 236 --LKWAPDEEGTIFLYKONGNEDRVTKAIEKISAKNKSOGRIEFPKPAIT--- 350  
DQ 237 INLKSIPKREKELIYLDDOKKFSERKYSGLSLKGLKSGJGRLDGFQVYKREKQ 356  
QY 351 -SAPLKRETSDDKTSKAAANKTK 373  
DQ 357 LAAAKRAQENKLNK-NKNKVK 379

RESULT 10

ABG19545  
ID ABG19545 standard; protein; 373 AA.  
XX  
AC ABG19545;  
XX  
DE 13-FEB-2002 (first entry)  
XX  
DE Novel human diagnostic protein #.9536.

Human: Chromosome mapping; gene mapping; gene therapy; forensic;  
KM Food supplement; medical imaging; diagnostic; genetic disorder.  
XX

OS Homo sapiens.

XX MO200275067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001MO-0508631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0549167.

XX (HVS-) HVSQ INC.

XX Dmanac RT, Liu C, Tang YF;

XX WPI: 2001-639362/73.

XX N-PSDB: AAS83732.

PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.

PS Claim 20: SEQ ID NO 49904; 103bp; English.

XX The invention relates to isolated polynucleotide (I) and  
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,  
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
XX and gene mapping, and in recombinant production of (II). The  
XX polynucleotides are also used in diagnostics as expressed sequence tags  
XX for identifying expressed genes. (I) is useful in gene therapy techniques  
XX to restore normal activity of (II) or to treat disease states involving  
XX (II). (II) is useful for generating antibodies against it, detecting or  
XX quantitating a polypeptide in tissue, as molecular weight markers and as  
XX a food supplement. (II) and its binding partners are useful in medical  
XX imaging of sites expressing (II). (I) and (II) are useful for treating  
XX disorders involving aberrant protein expression or biological activity.  
XX The polypeptide and polynucleotide sequences have applications in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits to assess biodiversity  
XX and to produce other types of data and products dependent on DNA and  
XX amino acid sequences. ABG00010-ABG30377 represent novel human  
XX diagnostic amino acid sequences of the invention.  
XX Note: The sequence data for this patent did not appear in the printed  
XX specification, but was obtained in electronic format directly from WIPQ  
XX at ftp.wipo.int/pub/published\_pcl\_sequences.

XX Sequence 373 AA;

Query Match 36.1%; Score 659.5; DB 22; Length: 373;

Best Local Similarity 41.1%; Pired. No. 5, 4e-54;  
Matches 155; Conservative 68; Mismatches 105; Indels 49; Gaps

QY 1 KGTICGLTKLLADNAPAMKQKKESTFGRRIVAVASMSIYCLIVGRQMETLINEAGE  
DQ 35 MGIOGLAKLLIADAPASIRENDIKSYGRKVAIDASMSIYFLIIV-ROGGVILONEGE  
QY 61 VTSHLQGMENRTIRLLEAGIKPYVVDGKIPDOKKQELAKRYSRDDATKDLTEAVEGD  
DQ 95 TSHLMQMYRTIRKFKNGCIKIPY-----KQGLAKCSFLKAFKFKLOOAOAGA  
QY 121 KDAIEKLSKRTVVTROHNECKRLILMLGVPEVAESEAARCAALCINDKFAVASED  
DQ 146 EGYEKETKVLVAVTKQHNCKCHLLSLMGIPLDAPLEAFACALVAKGKYPVATED  
QY 181 MDSLTGAPRFLNHLMPSSKRLPWEFVAKYLELELIMDQFIDLCILGCTNYDSIK  
DQ 206 MDCFTGSPVLMQHLTASS-----DYCKSIW  
QY 241 GIGGOTALKLIRQHSIFSLLENLNDKRVQIPEDMPEYQARLRFKPNVT-LDIPELKW  
DQ 232 SIGKRAVDILQHKSTEEVQKLDPRKPYLPNNRLHKEAYCFLFEFV-GPYSVILKWS  
QY 300 APPEEGTIFLYKONGNEDRVTKAIEKISAKNKSOGRIEFPKPAITTSAPLRKE  
DQ 292 EPNDEELVKFEQGEKOPSEERIRHSYVKRLKTEKG-GGEFPGPWIQP--GRASPIKRPDH  
QY 359 TSDDKTSKAAANKKTKAG 375  
DQ 349 QEDGHTSRSSDSGSHKG 365

RESULT 11

XX AAM24216

XX AAM24216 standard; protein; 340 AA.

XX AAM24216;

XX 25-MAR-2003 (updated)

XX 14-APR-1998 (first entry)

XX Pyrococcus furiosus FEN-1 endonuclease.

XX Nucleic acid cleavage; 5' nuclease; DNA cleavage; RNA cleavage;  
XX invader directed cleavage; FEN-1; endonuclease.

XX Pyrococcus furiosus.

XX MO9727214-A1.

XX 31-JUL-1997.

XX 22-JAN-1997; 97MO-0501072.

XX 02-DEC-1996; 96US-0759038.

XX 24-JAN-1996; 96US-0599491.

XX 12-JUL-1996; 96US-0682853.

XX 29-NOV-1996; 96US-0756386.

XX 02-DEC-1996; 96US-0758314.

XX (THIR-) THIRD WAVE TECHNOLOGIES INC.

XX Hall JG, Lyamichov VI, Prudent JR, Brow MWD, Kaiser MW;

XX Lyamichov N, Olive DM, Dahlberg JE;

XX WPI: 1997-39361/36.

XX N-PSDB: AAT76685.

XX Thermostable structure-specific nuclease(s) - used for detection and  
XX characterisation of nucleic acid sequences and variations in nucleic  
XX acid sequences

PS Example 28; Page 283-285; 457pp; English.

XX This sequence comprises *Pyrococcus furiosus* (Pfu) FEN-1  
 CC endonuclease. Large-scale production of the enzyme was performed  
 CC using *E. coli* host cells transformed with a vector carrying the Pfu  
 CC FEN-1 coding sequence (see AAT76685). Pfu FEN-1 is a thermostable  
 CC enzyme. It can be used in novel methods for the detection and  
 CC characterisation of nucleic acid sequences and variations in  
 CC nucleic acid sequences.  
 CC (updated on 25-MAR-2003 to correct PI field.)  
 CC  
 XX Sequence 340 AA;  
 SQ  
 Query Match 33.4%; Score 648.5; DB 18; Length 340;  
 Best Local Similarity 41.3%; Pred. No. 2,4e-51;  
 Matches 142; Conservative 62; Mismatches 107; Indels 33; Gaps 6;  
 QY 19 KEQKFESYFGKRIAVDASMSITVGLTGMETLTNEGVSHTLQGMNRTLLA 79  
 DB 12 KEIELENLYGKRIADALNAIYQFLSTIRKQDPTPLMSKGRITSHLSGLFYKINLMEA 71  
 QY 79 GIKPYVYVDPGPPDKKQELAKRYSKRDPATKDLTFAVEVGDKDAIEKLSKRYKTAGH 138  
 DB 72 GIKPYVYVDPGPPDKKQELAKRYSKRDPATKDLTFAVEVGDKDAIEKLSKRYKTAGH 138  
 QY 139 NEDCKRLRLMGVYVPAFSEAECAALCINDKFAVASEDMSLTFGAPRLRLHMDP 198  
 DB 132 IEDAKKLLLELMGIPYQAPSEGEAOAAVMKAGSYVASODYDLSLFGAPRLVRLNLTIT 191  
 QY 199 SSKKIP-----VMEFDVAKYLELELTMDQFIDLCIGCDY-CDSIKGIG 244  
 DB 192 GKRKLPGKKNVYEIKPELITLIE-----EVLKELKLTREKLELALVGTDPNPGIKIGL 247  
 QY 245 QTALKLIRHGSISTELENKND---RYQIPEDMPYOEARLRFEPNVTIDIPELKWTAP 301  
 DB 248 KKALEIVRH-----SKDPLAKFKOSQSDVLYAKKEFLNPVPT-DNYNLVWBP 295  
 QY 302 DEEGLISFLVKNQGNEDVYTKALEIKSAKNSKSGRLESPFK 345  
 DB 296 DEEGLKFLCDHDFSEFRVKNGLERLKAIKSGKSTLESWPK 339  
 RESULT 12  
 ID AAM79970 standard: Protein: 340 AA.  
 XX  
 AC AAM79970:  
 XX  
 DT 02-FEB-1999 (first entry)  
 XX  
 DE *Pyrococcus furiosus* FEN-1 endonuclease.  
 XX  
 KM Nucleic acid detection; multiple sequential invasive cleavage;  
 XX FEN-1; endonuclease; nuclease.  
 OS *Pyrococcus furiosus*.  
 XX  
 PN W09842873-A1.  
 XX  
 CD 01-OCT-1998.  
 XX  
 EF 24-MAR-1998; 3AWO-GS05809.  
 XX  
 PR 24-MAR-1997; 97US-0823516.  
 XX  
 PA (THIR-) THIRD WAVE TECHNOLOGIES INC.  
 XX  
 PI Brock MAC, Hall JG, Kwiatkowski RW, Lyamichew V;  
 PI Mast AL, Vavra SH;  
 XX WPI: 1998-557036/47.  
 CR N-PSDB; AAV65840.  
 XX  
 PT Detecting target nucleic acid by sequence-specific cleavage of

PT complex with two specific oligonucleotides - used to detect  
 PT cytomegalovirus DNA  
 XX  
 PS Example 28b: Page 316-317; 52app; English.  
 XX  
 CC This is the amino acid sequence of FEN-1 endonuclease of  
 CC *Pyrococcus furiosus* (Pfu). FEN-1 DNA (see AAV65840) has been  
 CC ligated into vector for expression in *E. coli* cells. The invention  
 CC relates to means for the detection and characterisation of nucleic  
 CC acid sequences, and variations in nucleic acid sequences. It also  
 CC relates to methods for forming a nucleic acid cleavage structure on  
 CC a target sequence and cleaving this structure in a site-specific  
 CC manner, preferably using a thermostable structure-specific nuclease  
 CC such as FEN-1. Cleavage of the cleavage structure by the  
 CC nuclease indicates the presence of specific nucleic acid sequences  
 CC or specific variants. The invention further relates to methods for  
 CC the separation of nucleic acid molecules based on charge, methods  
 CC for the detection of non-target cleavage products via the formation  
 CC of a complete and activated protein binding region, and methods for  
 CC the detection of nucleic acid from various viruses (e.g. human  
 CC cytomegalovirus) in a sample. The method amplifies the detection  
 CC molecule rather than the target itself, is less subject to  
 CC contamination than exponential amplification processes, and allows  
 CC many targets to be analysed in a single reaction.  
 CC  
 XX Sequence 340 AA;  
 SQ  
 Query Match 33.4%; Score 648.5; DB 19; Length 340;  
 Best Local Similarity 41.3%; Pred. No. 2,4e-51;  
 Matches 142; Conservative 62; Mismatches 107; Indels 33; Gaps 6;  
 QY 19 KEQKFESYFGKRIAVDASMSITVGLTGMETLTNEGVSHTLQGMNRTLLA 79  
 DB 12 KEIELENLYGKRIADALNAIYQFLSTIRKQDPTPLMSKGRITSHLSGLFYKINLMEA 71  
 QY 79 GIKPYVYVDPGPPDKKQELAKRYSKRDPATKDLTFAVEVGDKDAIEKLSKRYKTAGH 138  
 DB 72 GIKPYVYVDPGPPDKKQELAKRYSKRDPATKDLTFAVEVGDKDAIEKLSKRYKTAGH 138  
 QY 139 NEDCKRLRLMGVYVPAFSEAECAALCINDKFAVASEDMSLTFGAPRLRLHMDP 198  
 DB 132 IEDAKKLLLELMGIPYQAPSEGEAOAAVMKAGSYVASODYDLSLFGAPRLVRLNLTIT 191  
 QY 199 SSKKIP-----VMEFDVAKYLELELTMDQFIDLCIGCDY-CDSIKGIG 244  
 DB 192 GKRKLPGKKNVYEIKPELITLIE-----EVLKELKLTREKLELALVGTDPNPGIKIGL 247  
 QY 245 QTALKLIRHGSISTELENKND---RYQIPEDMPYOEARLRFEPNVTIDIPELKWTAP 301  
 DB 248 KKALEIVRH-----SKDPLAKFKOSQSDVLYAKKEFLNPVPT-DNYNLVWBP 295  
 QY 302 DEEGLISFLVKNQGNEDVYTKALEIKSAKNSKSGRLESPFK 345  
 DB 296 DEEGLKFLCDHDFSEFRVKNGLERLKAIKSGKSTLESWPK 339  
 RESULT 12  
 ID AAM59940 standard: Protein: 340 AA.  
 XX  
 AC AAM59940:  
 XX  
 DT 21-DEC-1998 (first entry)  
 XX  
 DE Amino acid sequence of the Pfu FEN-1 endonuclease.  
 XX  
 KM Mja FEN-1 endonuclease; Tag gene; structure-specific nuclease;  
 KM mutant; DNA polymerase; bacterial; fungi; protozoa; RNA virus;  
 XX hepatitis C virus; HIV; thermostable.  
 OS *Pyrococcus furiosus*.  
 XX  
 PN W09821774-A1.



XX 04-JUN-1998.  
 PD  
 XX 26-NOV-1997; 97WD-US21783.  
 PF  
 XX 02-DEC-1996; 96US-0758314.  
 PR  
 XX 29-NOV-1996; 96US-0757653.  
 XX  
 PA (THIR-) THIRD WAVE TECHNOLOGIES INC.  
 XX  
 PI Kaiser MW, Lyamichev VI, Lyamicheva N;  
 DR MPI: 1998-322748/28.  
 DR N-PSDB: AAV53951.  
 XX  
 PT Thermostable structure-specific nuclease(s) derived from mutant DNA  
 PT polymerase(s) - useful for detecting mutant allele(s) or strains of  
 PT microorganisms  
 XX  
 PS Example 29; Pages 280-281; 472pp; English.  
 XX  
 CC This is the amino acid sequence encoding the Pfu FEN-1 endonuclease,  
 CC used in the method of the invention. In this process thermostable  
 CC structure-specific nucleases are derived from mutant DNA polymerases,  
 CC which can be used for detecting mutant alleles or strains of  
 CC microorganisms. The structure-specific nucleases can be used in  
 CC mixtures, compositions and kits to treat nucleic acid, e.g. for  
 CC detection of wild type and mutant alleles of genes, for detection  
 CC and/or identification of strains of microorganisms such as bacteria,  
 CC fungi, protozoa, especially for detection of RNA viruses such as the  
 CC hepatitis C virus (HCV).  
 CC  
 SQ Sequence 340 AA:  
 Query Match 33.4%; Score 648.5; DB 19; Length 340;  
 Best Local Similarity 41.3%; Pred. No. 2,4e-51;  
 Matches 142; Conservative 62; Mismatches 107; Indels 33; Gaps 6;  
 QY 19 KECKFESYFGRIKAVDASHSIYQFLIVGRIGMETLTNEAGEVTSHQGMNPTITLLEA 78  
 DB 12 KEILENLVGGKIALDALNLYGFLSTIRQKDGITPLMSKGRITSHLSGLFRTINLMEA 71  
 QY 79 GIKPVYVFDGKPPDKKQQLAKRYSKRDPATKDLTEAVEVGDKDAIEKLSKRTYKVTROH 138  
 DB 72 GIKPVYVFDGEPPEPKKKKELEKREAREAEKKEALEKEIDEARKYACRATRVNEML 131  
 QY 139 NEDCKRLRLMGVPPVPAPEAEACALCINDKFAVASDMDSLTGAPRLRLHLMKP 198  
 DB 132 IEDAKRLLELWGIPIVOAPSEGAQAAVMAAKGSYASASQYOSLLFGAPRLVRNLTIT 191  
 QY 199 SSKRIP-----VMEFVAKVLELELTMDQFIDLCITGCDY-CDSITKQIGG 244  
 DB 192 GKRLPGKKNVVEIKPELTILE-----EVLKELKLTREKLTIELALVGDYMDGSKGICG 247  
 QY 245 QTAALKLIRQHGSIESTLENLND---RYQIPEDWPYQEARLREKPNVTLIDIPHLKWTAP 301  
 DB 248 KKALEIVRH-----SKDPLAKFQKQSDVDLVAIKKEFLNPVPT-DNYNLVWRDP 295  
 QY 302 DEGLISFLVKNQGFNEDRYTKAIEKIKSAKSKSSQGLSESFFK 345  
 DB 296 DEGLIKFLCDHDFSEERVKNGLERLKAIKSGKOSTLESFVK 339  
 RESULT 14  
 ABG92941  
 ID ABG92941 standard; Protein; 340 AA.  
 XX  
 AC ABG92941:  
 XX  
 XX 20-NOV-2002 (first entry)  
 XX  
 DE Pyrococcus furiosus FEN1 protein.

KM Phosphoramidite; INVADER assay cleavage reaction; FEN1; cleavage;  
 KM nucleic acid separation; DNA polymerase; human; MCF-1; ubiquitin;  
 KM monocytic chemottractant protein-1; enzyme.  
 OS  
 XX Pyrococcus furiosus.  
 XX  
 PN W020026304C-A2.  
 XX  
 PD 15-ACG-2002.  
 XX  
 PF 06-FEB-2002; 2002WD-US03423.  
 XX  
 PR 06-FEB-2001; 2001CS-0777430.  
 XX  
 PA (THIR-) THIRD WAVE TECHNOLOGIES INC.  
 XX  
 PI Lyamichev V, Skrzeczynski Z, Allawi HT, Mayland SR, Takova T;  
 PI Neri BP;  
 XX  
 DR MPI: 2002-674850/72.  
 XX  
 PT Composition useful for e.g. separation of nucleic acids comprises a  
 PT positively or neutrally charged phosphoramidite -  
 XX  
 PS Example 2; Page 179-180; 197pp; English.  
 XX  
 CC The invention relates to a composition comprising a positively or  
 CC neutrally charged phosphoramidite. The composition is useful for  
 CC separation of nucleic acid molecules. The composition is further useful  
 CC for fractionation of specific nucleic acids by selective charge revers  
 CC useful in e.g. INVADER assay cleavage reactions; and in the synthesis  
 CC of charge-balanced molecules. In the fractionation of nucleic acid  
 CC molecules, the method provides an absolute readout of the partition  
 CC products from substrates (i.e. provides a 100% separation). Through th  
 CC use of multiple positively charged adducts, synthetic molecules can be  
 CC constructed with sufficient solubilization due to the fact that the  
 CC normally negatively charged strand is made nearly neutral. It is also  
 CC possible to distinguish between an enzymatically or thermally degraded  
 CC DNA fragments due to the absence or presence of 3' phosphate.  
 CC  
 SQ Sequence 340 AA:  
 Query Match 33.4%; Score 648.5; DB 23; Length 340;  
 Best Local Similarity 41.3%; Pred. No. 2,4e-51;  
 Matches 142; Conservative 62; Mismatches 107; Indels 33; Gaps 6;  
 QY 19 KECKFESYFGRIKAVDASHSIYQFLIVGRIGMETLTNEAGEVTSHQGMNPTITLLEA 78  
 DB 22 KEILENLVGGKIALDALNLYGFLSTIRQKDGITPLMSKGRITSHLSGLFRTINLMEA 71  
 QY 79 GIKPVYVFDGKPPDKKQQLAKRYSKRDPATKDLTEAVEVGDKDAIEKLSKRTYKVTROH 138  
 DB 72 GIKPVYVFDGEPPEPKKKKELEKREAREAEKKEALEKEIDEARKYACRATRVNEML 131  
 QY 139 NEDCKRLRLMGVPPVPAPEAEACALCINDKFAVASDMDSLTGAPRLRLHLMKP 198  
 DB 132 IEDAKRLLELWGIPIVOAPSEGAQAAVMAAKGSYASASQYOSLLFGAPRLVRNLTIT 191  
 QY 199 SSKRIP-----VMEFVAKVLELELTMDQFIDLCITGCDY-CDSITGIGIG 244  
 DB 192 GKRLPGKKNVVEIKPELTILE-----EVLKELKLTREKLTIELALVGDYMDGSKGICG 247  
 QY 245 QTAALKLIRQHGSIESTLENLND---RYQIPEDWPYQEARLREKPNVTLIDIPHLKWTAP 301  
 DB 248 KKALEIVRH-----SKDPLAKFQKQSDVDLVAIKKEFLNPVPT-DNYNLVWRDP 295  
 QY 302 DEGLISFLVKNQGFNEDRYTKAIEKIKSAKSKSSQGLSESFFK 345  
 DB 296 DEGLIKFLCDHDFSEERVKNGLERLKAIKSGKOSTLESFVK 339  
 RESULT 15





## RESULT 2

A56531

DNA structure-specific endonuclease FEN1 [validated] - human

K:Alternate names: flap endonuclease 1; rad2 protein homolog

C:Species: Homo sapiens (man)

C:Date: 19-Oct-1995 #sequence\_revision 19-Oct-1995 #text\_change 21-Jul-2000

C:Accession: A56531; B56054

R:Hirooka, T.R.; Harrington, J.J.; Gerhard, D.S.; Lieber, M.R.; Hsieh, C.L.

Genomics 25: 220-225, 1995

A&gt;Title: Sequence of human FEN-1, a structure-specific endonuclease and chromosomal loop

A:Reference number: A56531; M010:95253376; PMID:7774922

A:Accession: A56531

A&gt;Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-380 &lt;Hir&gt;

A:Cross-references: GB:L37374; NID:q642089; PUD:AAA91331.1; PID:q642089

R:Murphy, J.M.; Tavassoli, M.; Al-Harithy, R.; Sheldrick, K.S.; Lehmann, A.R.; Carr, A.M.

Mol. Cell. Biol. 14: 4878-4888, 1994

A&gt;Title: Structural and functional conservation of the human homolog of the Schizosaccha

A:Reference number: A56054; M010:94277093; PMID:8007985

A:Accession: B56054

A&gt;Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-380 &lt;MUR&gt;

A:Cross-references: GB:S70527

C:Genetics:

A:Gene: GDB:FEN1

A:Cross-references: GDB:451178; OMIM:600393

A:Map position: 11q12-11q12

C:Function:

A:Description: cleaves substrates with unannealed 5'-tails [validated, M010:20204428]

A&gt;Note: needed during Okazaki fragment processing and long patch base excision repair, C

C:Superfamily: conserved hypothetical protein YK113C

C:Keywords: DNA repair; endonuclease

Query Match

Best Local Similarity 53.8%; Pred. No. 1e-61; Length 380;

Matches 206; Conservative 69; Mismatches 98; Indels 10; Gaps 6;

QY 1 MGIGKLTLLADNAPKAMKQKFESEYFGKRIADVASMSTYQFLVVGRTGMETLTNEAGE 60

DB 1 MGIGKLTLLADNAPKAMKQKFESEYFGKRIADVASMSTYQFLVVGRTGMETLTNEAGE 60

DB 1 MGIGKLTLLADNAPKAMKQKFESEYFGKRIADVASMSTYQFLVVGRTGMETLTNEAGE 60

QY 61 VSHLQGMFNRTIRLLLEAGIKPVYVFDGKPPDMKKKQELAKRYSKRDAITKLTAVAVGD 120

DB 61 VSHLQGMFNRTIRLLLEAGIKPVYVFDGKPPDMKKKQELAKRYSKRDAITKLTAVAVGD 120

DB 61 VSHLQGMFNRTIRLLLEAGIKPVYVFDGKPPDMKKKQELAKRYSKRDAITKLTAVAVGD 120

QY 62 TSHLMGMFNRTIRLLLEAGIKPVYVFDGKPPDMKKKQELAKRYSKRDAITKLTAVAVGD 119

DB 62 TSHLMGMFNRTIRLLLEAGIKPVYVFDGKPPDMKKKQELAKRYSKRDAITKLTAVAVGD 119

DB 62 TSHLMGMFNRTIRLLLEAGIKPVYVFDGKPPDMKKKQELAKRYSKRDAITKLTAVAVGD 119

QY 121 KDALEKSKRTVYKVRQHNEDCKRLRLMGVYVFAFSEAECAALTCINCFVAVASD 180

DB 121 KDALEKSKRTVYKVRQHNEDCKRLRLMGVYVFAFSEAECAALTCINCFVAVASD 180

DB 121 KDALEKSKRTVYKVRQHNEDCKRLRLMGVYVFAFSEAECAALTCINCFVAVASD 180

QY 120 EGVVEKFKRLKLVTKQHNDCKRLRLMGVYVFAFSEAECAALTCINCFVAVASD 179

DB 120 EGVVEKFKRLKLVTKQHNDCKRLRLMGVYVFAFSEAECAALTCINCFVAVASD 179

DB 120 EGVVEKFKRLKLVTKQHNDCKRLRLMGVYVFAFSEAECAALTCINCFVAVASD 179

QY 181 MDSLTFGAPRFRHLMDPSSKIPVMEFVAVKLELELIMDQFLDCLCGCCYCDISK 240

DB 181 MDSLTFGAPRFRHLMDPSSKIPVMEFVAVKLELELIMDQFLDCLCGCCYCDISK 240

DB 181 MDSLTFGAPRFRHLMDPSSKIPVMEFVAVKLELELIMDQFLDCLCGCCYCDISK 240

QY 130 MDCLTFFGSPVLRHFLASEAKKLPFOEHLSPRLQELNCEQFVDCILLSDYCESIR 239

DB 130 MDCLTFFGSPVLRHFLASEAKKLPFOEHLSPRLQELNCEQFVDCILLSDYCESIR 239

DB 130 MDCLTFFGSPVLRHFLASEAKKLPFOEHLSPRLQELNCEQFVDCILLSDYCESIR 239

QY 241 GIGGCTALKLIRHOGSISTELENKDRYOIPEDMPYQEARLFEKPEVITDIP-----E 298

DB 241 GIGGCTALKLIRHOGSISTELENKDRYOIPEDMPYQEARLFEKPEVITDIP-----E 298

DB 241 GIGGCTALKLIRHOGSISTELENKDRYOIPEDMPYQEARLFEKPEVITDIP-----E 298

QY 240 GIGGCTALKLIRHOGSISTELENKDRYOIPEDMPYQEARLFEKPEVITDIP-----E 298

DB 240 GIGGCTALKLIRHOGSISTELENKDRYOIPEDMPYQEARLFEKPEVITDIP-----E 298

DB 240 GIGGCTALKLIRHOGSISTELENKDRYOIPEDMPYQEARLFEKPEVITDIP-----E 298

QY 239 TAVDEGLISFLVKNQNGNEDRYTKAEIKSAKNSQGRLESFFKPTACTSAFKRKE 354

DB 239 TAVDEGLISFLVKNQNGNEDRYTKAEIKSAKNSQGRLESFFKPTACTSAFKRKE 354

DB 239 TAVDEGLISFLVKNQNGNEDRYTKAEIKSAKNSQGRLESFFKPTACTSAFKRKE 354

QY 230 SEPNEEELIKTMCGKOPSEERIRSGYRLSKRSQSGTQGRLDFFKVTGLSS-AKRKE 357

DB 230 SEPNEEELIKTMCGKOPSEERIRSGYRLSKRSQSGTQGRLDFFKVTGLSS-AKRKE 357

DB 230 SEPNEEELIKTMCGKOPSEERIRSGYRLSKRSQSGTQGRLDFFKVTGLSS-AKRKE 357

QY 359 TSDK--TSKAANKKTKAGKKK 379

DB 359 TSDK--TSKAANKKTKAGKKK 379

DB 359 TSDK--TSKAANKKTKAGKKK 379

QY 358 PEKGSYKKKA---KTGANGKFK 377

DB 358 PEKGSYKKKA---KTGANGKFK 377

DB 358 PEKGSYKKKA---KTGANGKFK 377

C:Accession: A56054; T38725

R:Murphy, J.M.; Tavassoli, M.; Al-Harithy, R.; Sheldrick, K.S.; Lehmann, A.

Mol. Cell. Biol. 14: 4878-4888, 1994

A&gt;Title: Structural and functional conservation of the human homolog of the

A:Reference number: A56054; M010:94277093; PMID:8007985

A:Accession: A56054

A&gt;Status: preliminary

A:Molecule type: DNA

A:Residues: 1-380 &lt;MUR&gt;

A:Cross-references: GB:S70605

A:Accession: T45210

A&gt;Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-380 &lt;MUR&gt;

A:Cross-references: EMBL:X73011; NID:q2909339; PUD:CA8499.1; PID:q445685

A:Experimental source: strain 972H(-)

R:Genetics, S.; Churcher, C.M.; Barrett, B.C.; Rajandream, M.A.; Wood, V.

submitted to the EMBL Data Library, September 1997

A:Reference number: 221797

A:Accession: T38725

A&gt;Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-380 &lt;GEN&gt;

A:Cross-references: EMBL:299167; PUD:CAP1632.1; GSPDB:GNC0066; SPDB:SPAC3

A:Experimental source: strain 972H(-) cosmid c386

C:Genetics:

A:Gene: rad2

A:Map position: 1

A:Functions: 5/1; 35/2

C:Function:

A:Description: DNA repair

C:Superfamily: conserved hypothetical protein YK113C

Query Match

Best Local Similarity 50.7%; Pred. No. 3.8e-61; Length 380;

Matches 194; Conservative 73; Mismatches 107; Indels 9; Gaps

QY 1 MGIGKLTLLADNAPKAMKQKFESEYFGKRIADVASMSTYQFLVVGRTGMETLTNEAGE 60

DB 1 MGIGKLTLLADNAPKAMKQKFESEYFGKRIADVASMSTYQFLVVGRTGMETLTNEAGE 60

DB 1 MGIGKLTLLADNAPKAMKQKFESEYFGKRIADVASMSTYQFLVVGRTGMETLTNEAGE 60

QY 61 VSHLQGMFNRTIRLLLEAGIKPVYVFDGKPPDMKKKQELAKRYSKRDAITKLTAVAVGD 120

DB 61 VSHLQGMFNRTIRLLLEAGIKPVYVFDGKPPDMKKKQELAKRYSKRDAITKLTAVAVGD 120

DB 61 VSHLQGMFNRTIRLLLEAGIKPVYVFDGKPPDMKKKQELAKRYSKRDAITKLTAVAVGD 120

QY 62 TSHLMGMFNRTIRLLLEAGIKPVYVFDGKPPDMKKKQELAKRYSKRDAITKLTAVAVGD 119

DB 62 TSHLMGMFNRTIRLLLEAGIKPVYVFDGKPPDMKKKQELAKRYSKRDAITKLTAVAVGD 119

DB 62 TSHLMGMFNRTIRLLLEAGIKPVYVFDGKPPDMKKKQELAKRYSKRDAITKLTAVAVGD 119

QY 121 KDALEKSKRTVYKVRQHNEDCKRLRLMGVYVFAFSEAECAALTCINCFVAVASD 180

DB 121 KDALEKSKRTVYKVRQHNEDCKRLRLMGVYVFAFSEAECAALTCINCFVAVASD 180

DB 121 KDALEKSKRTVYKVRQHNEDCKRLRLMGVYVFAFSEAECAALTCINCFVAVASD 180

QY 120 AEMVSEFAKRYKVRQHNEDCKRLRLMGVYVFAFSEAECAALTCINCFVAVASD 179

DB 120 AEMVSEFAKRYKVRQHNEDCKRLRLMGVYVFAFSEAECAALTCINCFVAVASD 179

DB 120 AEMVSEFAKRYKVRQHNEDCKRLRLMGVYVFAFSEAECAALTCINCFVAVASD 179

QY 181 MDSLTFGAPRFRHLMDPSSKIPVMEFVAVKLELELIMDQFLDCLCGCCYCDISK 240

DB 181 MDSLTFGAPRFRHLMDPSSKIPVMEFVAVKLELELIMDQFLDCLCGCCYCDISK 240

DB 181 MDSLTFGAPRFRHLMDPSSKIPVMEFVAVKLELELIMDQFLDCLCGCCYCDISK 240

QY 242 GIGGCTALKLIRHOGSISTELENKDRYOIPEDMPYQEARLFEKPEVITDIP-----E 298

DB 242 GIGGCTALKLIRHOGSISTELENKDRYOIPEDMPYQEARLFEKPEVITDIP-----E 298

DB 242 GIGGCTALKLIRHOGSISTELENKDRYOIPEDMPYQEARLFEKPEVITDIP-----E 298

QY 241 GAVPAFAVELIRYQGLDFVFNACSKYPIEDMPYQEARLFEKPEVITDIP-----E 298

DB 241 GAVPAFAVELIRYQGLDFVFNACSKYPIEDMPYQEARLFEKPEVITDIP-----E 298

DB 241 GAVPAFAVELIRYQGLDFVFNACSKYPIEDMPYQEARLFEKPEVITDIP-----E 298

QY 297 KWIAPDEGLISFLVKNQNGNEDRYTKAEIKSAKNSQGRLESFFKPTACTSAFKRKE 354

DB 297 KWIAPDEGLISFLVKNQNGNEDRYTKAEIKSAKNSQGRLESFFKPTACTSAFKRKE 354

DB 297 KWIAPDEGLISFLVKNQNGNEDRYTKAEIKSAKNSQGRLESFFKPTACTSAFKRKE 354

QY 298 KKSPPDGLISFLVKNQNGNEDRYTKAEIKSAKNSQGRLESFFKPTACTSAFKRKE 357

DB 298 KKSPPDGLISFLVKNQNGNEDRYTKAEIKSAKNSQGRLESFFKPTACTSAFKRKE 357

DB 298 KKSPPDGLISFLVKNQNGNEDRYTKAEIKSAKNSQGRLESFFKPTACTSAFKRKE 357

QY 357 KERSDUTSKAANKKTKAGKKK 379

DB 357 KERSDUTSKAANKKTKAGKKK 379

DB 357 KERSDUTSKAANKKTKAGKKK 379

QY 356 VIKRSGSKAKKKKSKGSKK 478

DB 356 VIKRSGSKAKKKKSKGSKK 478

DB 356 VIKRSGSKAKKKKSKGSKK 478

QY 357 KERSDUTSKAANKKTKAGKKK 379

DB 357 KERSDUTSKAANKKTKAGKKK 379

DB 357 KERSDUTSKAANKKTKAGKKK 379

QY 356 VIKRSGSKAKKKKSKGSKK 478

DB 356 VIKRSGSKAKKKKSKGSKK 478

DB 356 VIKRSGSKAKKKKSKGSKK 478

## RESULT 4

A52730

DNA endonuclease 1 - mouse

K:Alternate names: FEN-1

C:Species: Mus musculus (house mouse)

C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 24-Sep-1999

C:Accession: A53730

R:Harrington, J.J.; Lieber, M.R.

Genes Dev. 8, 1344-1355, 1994

A>Title: Functional domains within FEN-1 and RAD2 define a family of structure-specific

A:Reference number: A53730; M01D:95011546; PMID:7926735

A:Accession: A53730

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-378 <HAR>

A:Cross-references: GB:U126320; NID:q499648; P1DN:CAA37664.1; P1D:q499649

C:Superfamily: conserved hypothetical protein YK113c

C:Keywords: DNA repair; endonuclease

Query Match

52.3%; Score 1014; DB 2; Length 378;

Best Local Similarity 53.2%; Pred. No. 1,6e-60;

Matches 202; Conservative 69; Mismatches 101; Indels 8; Gaps 6;

QY 1 MGKIGITKLADNNAKAKKEQKFESEYFGKRIAVDASMSIYOFLIVYGRTGHTLTNEAGE 60

DB 1 MGHIGIAKLADVAASAIKRENDIKSYFGKRIAVDASMSIYOFLIAY-FOGSDVLNENGE 59

QY 61 VSHLQGMENRTIRLEAGIKPYVYEDGKPPDMKQOELAKRYSKRDATXDLTAVEVGD 120

DB 60 TTS-LMGWYRTIR-MENKIRPVYVFDGKPPOLKSGELAKRSRAEKEKQLOQNOEAGM 127

QY 121 KDAIEKSKRTVAVTROHNECDKRLRLMGVPEVPEASEAEACALCINDKVFVASED 180

DB 119 EREVERKTRIVKVKVKNHDECKHLLMGHLYDPAPEAFASCAIAKAGVVAATED 177

QY 181 MDSITFGAPRFRLHLMDSKSKIPVMEFDVAKVPELTLTMDQFDLCILGGQVCSIR 240

DB 178 MDCLIFGSPVLMRHLTASAKK-LPIQERHLSKVLDELGINOCFVDLCLLGSNDCEIR 237

QY 241 GIGGQATALKLRQSGSIESILENLKDRYOIPEDMFOEARLFKEPNVT-ID-PELKMT 299

DB 236 GIGAKRAVDLIQKHSEIEIVRLDPSKYVPEENMLHREAQQLFEZVVCPESEYELKMS 297

QY 300 APDEGLISFLVKNQCFNEDRYTKAIEKISAKKNSOGRCJSPFKPATTISAPLKSET 359

DB 298 EPNEELVFMCGEKFSEIRISGVKRLSKSRGOSTGR-IDPEFKVYGSJSS-AKREKP 356

QY 360 SDPKSKAANKKTKAGGKK 379

DB 357 E---FKGPAKKKAKTGCGAK 373

QY 357 E---FKGPAKKKAKTGCGAK 373

DB 357 E---FKGPAKKKAKTGCGAK 373

QY 357 E---FKGPAKKKAKTGCGAK 373

DB 357 E---FKGPAKKKAKTGCGAK 373

QY 357 E---FKGPAKKKAKTGCGAK 373

DB 357 E---FKGPAKKKAKTGCGAK 373

QY 357 E---FKGPAKKKAKTGCGAK 373

DB 357 E---FKGPAKKKAKTGCGAK 373

QY 357 E---FKGPAKKKAKTGCGAK 373

DB 357 E---FKGPAKKKAKTGCGAK 373

QY 357 E---FKGPAKKKAKTGCGAK 373

DB 357 E---FKGPAKKKAKTGCGAK 373

QY 357 E---FKGPAKKKAKTGCGAK 373

DB 357 E---FKGPAKKKAKTGCGAK 373

QY 357 E---FKGPAKKKAKTGCGAK 373

DB 357 E---FKGPAKKKAKTGCGAK 373

QY 357 E---FKGPAKKKAKTGCGAK 373

DB 357 E---FKGPAKKKAKTGCGAK 373

QY 357 E---FKGPAKKKAKTGCGAK 373

DB 357 E---FKGPAKKKAKTGCGAK 373

QY 357 E---FKGPAKKKAKTGCGAK 373

DB 357 E---FKGPAKKKAKTGCGAK 373

QY 357 E---FKGPAKKKAKTGCGAK 373

DB 357 E---FKGPAKKKAKTGCGAK 373

QY 357 E---FKGPAKKKAKTGCGAK 373

DB 357 E---FKGPAKKKAKTGCGAK 373

QY 61 VSHLQGMENRTIRLEAGIKPYVYEDGKPPDMKQOELAKRYSKRDATKDLTAVEVGD

DB 60 TTS-LMGWYRTIR-MENKIRPVYVFDGKPPOLKSGELAKRSRAEKEKQLOQNOEAGM

QY 121 KDAIEKSKRTVAVTROHNECDKRLRLMGVPEVPEASEAEACALCINDKVFVASED

DB 119 EREVERKTRIVKVKVKNHDECKHLLMGHLYDPAPEAFASCAIAKAGVVAATED

QY 181 MDSITFGAPRFRLHLMDSKSKIPVMEFDVAKVPELTLTMDQFDLCILGGQVCSIR

DB 178 MDCLIFGSPVLMRHLTASAKK-LPIQERHLSKVLDELGINOCFVDLCLLGSNDCEIR

QY 241 GIGGQATALKLRQSGSIESILENLKDRYOIPEDMFOEARLFKEPNVT-ID-PELKMT

DB 236 GIGAKRAVDLIQKHSEIEIVRLDPSKYVPEENMLHREAQQLFEZVVCPESEYELKMS

QY 300 APDEGLISFLVKNQCFNEDRYTKAIEKISAKKNSOGRCJSPFKPATTISAPLKSET

DB 298 EPNEELVFMCGEKFSEIRISGVKRLSKSRGOSTGR-IDPEFKVYGSJSS-AKREKP

QY 360 SDPKSKAANKKTKAGGKK 379

DB 357 E---FKGPAKKKAKTGCGAK 373

QY 357 E---FKGPAKKKAKTGCGAK 373

DB 357 E---FKGPAKKKAKTGCGAK 373

QY 357 E---FKGPAKKKAKTGCGAK 373

DB 357 E---FKGPAKKKAKTGCGAK 373

QY 357 E---FKGPAKKKAKTGCGAK 373

DB 357 E---FKGPAKKKAKTGCGAK 373

QY 357 E---FKGPAKKKAKTGCGAK 373

DB 357 E---FKGPAKKKAKTGCGAK 373

QY 357 E---FKGPAKKKAKTGCGAK 373

DB 357 E---FKGPAKKKAKTGCGAK 373

QY 357 E---FKGPAKKKAKTGCGAK 373

DB 357 E---FKGPAKKKAKTGCGAK 373

QY 357 E---FKGPAKKKAKTGCGAK 373

DB 357 E---FKGPAKKKAKTGCGAK 373

QY 357 E---FKGPAKKKAKTGCGAK 373

DB 357 E---FKGPAKKKAKTGCGAK 373

QY 357 E---FKGPAKKKAKTGCGAK 373

DB 357 E---FKGPAKKKAKTGCGAK 373

QY 357 E---FKGPAKKKAKTGCGAK 373

DB 357 E---FKGPAKKKAKTGCGAK 373

QY 357 E---FKGPAKKKAKTGCGAK 373

DB 357 E---FKGPAKKKAKTGCGAK 373

QY 357 E---FKGPAKKKAKTGCGAK 373

DB 357 E---FKGPAKKKAKTGCGAK 373

QY 357 E---FKGPAKKKAKTGCGAK 373

DB 357 E---FKGPAKKKAKTGCGAK 373

QY 357 E---FKGPAKKKAKTGCGAK 373

DB 357 E---FKGPAKKKAKTGCGAK 373

QY 357 E---FKGPAKKKAKTGCGAK 373

DB 357 E---FKGPAKKKAKTGCGAK 373

QY 357 E---FKGPAKKKAKTGCGAK 373

DB 357 E---FKGPAKKKAKTGCGAK 373

QY 357 E---FKGPAKKKAKTGCGAK 373

DB 357 E---FKGPAKKKAKTGCGAK 373

QY 357 E---FKGPAKKKAKTGCGAK 373

DB 357 E---FKGPAKKKAKTGCGAK 373

QY 357 E---FKGPAKKKAKTGCGAK 373

DB 357 E---FKGPAKKKAKTGCGAK 373

QY 181 MDLSTFCAPRLHIMDPSSKIPVMEFYAKVLEFLTMDFDICTCCGDDNSIK 243  
 Db 173 MDLSTFCAPRLHIMDPSSKIPVMEFYAKVLEFLTMDFDICTCCGDDNSIK 237  
 QY 241 GIGGOTALKLIRHOGSIESILEN-----NKRVOIPEDPYQEARLCKEPVWLTDFE 235  
 Db 238 GVGVTATKLIRHOGSIESILEN-----NKRVOIPEDPYQEARLCKEPVWLTDFE 236  
 QY 236 --LKWVAPDEGLISFLVKGNGNEDRVTKAIEKTSANKSSQGRLESFEPYAT 350  
 Db 297 INLKWSPKKEKEILEYCDCKKSEERVKSGISRLKGLKSGIOGRLDGPFVVPKIFQ 356  
 QY 353 --SAPLKRKETSDDTKSAANKKTK 373  
 Db 357 LAAAKRAOENKRLNK-NKKVTK 379

# RESULT 7

146893  
 endo/exonuclease fen-1 [imported] - Pyrococcus furiosus  
 N:Alternate names: DNA repair protein  
 C:Species: Pyrococcus furiosus  
 C:Date: 17-Mar-2000 \*sequence\_revision 17-Mar-2000 \*text\_change 21-Jul-2000  
 C:Accession: 146893  
 R:Disruptor: J. Brown, J. R. Bogert, A. P. Robb, F. T.  
 J. Mol. Evol. 49, 474-484, 1999  
 A>Title: DNA repair systems in Archaea: memories from the last universal common ancestor  
 A:Reference number: 224126; MUID:994.5851; PMID:10486005  
 A:Accession: 146893  
 A>Status: preliminary; translated from GS/FMBL/DB8  
 A:Molecule type: DNA  
 A:Residues: 1-340 <DIR>  
 A:Cross-references: EMBL:AF013497; NID:q4102601; PID:AA001514.1; PID:q4102602  
 A:Experimental source: strain DSM 3638  
 C:Genetics:  
 A:Gene: fen-1  
 C:Superfamily: conserved hypothetical protein YKL113c

Query Match 33.4%; Score 648.5; Db 2; Length 340;  
 Best Local Similarity 41.3%; Pred. No. 4.1e-36;  
 Matches 142; Conservative 62; Mismatches 107; Indels 33; Gaps 6;

QY 19 KEQKFEYFGKRIADVAMS.YQPIVAVGTGNETLTNEAGVYTHSGMNTTILFLA 78  
 Db 12 KEILENLYGKRIADVAMS.YQPIVAVGTGNETLTNEAGVYTHSGMNTTILFLA 71  
 QY 79 GIKPYYVDGKPPDMKQOELAKRYSKRDAKLTLEAVEGDKDAIEKLSKRTVYTRQ 138  
 Db 72 GIKPYYVDGKPPDMKQOELAKRYSKRDAKLTLEAVEGDKDAIEKLSKRTVYTRQ 131  
 QY 139 NEDCKRLRLMGVYVVEAPSAEACALCTINKVAVASDDSLTEGAPRCLRLMP 198  
 Db 132 EDCAKLELMGIPVAPSEGAAYMAKGSVASASODVDS.JIFGPRVIRN.TIT 193  
 QY 199 SSKRIPE-----WFEIVAAVLEELTLMQPIIDLCICGCV-CDISK3ICG 244  
 Db 192 GKRLPGKNYVYKPELLE-----EVKELKTRKTELIMAVIGTDPNGIGIGL 247  
 QY 245 QIALKLIRHOGSIESILEN.LNKD--RYQIPEDAPYQEARLCKEPVWLTDFE 301  
 Db 248 KRALHIVRH-----SKDPLAKFOKSDVDLVAIKFEFLNPPV.TENYVAVKRP 295  
 QY 302 DEGLISFLVKGNGNEDRVTKAIEKTSANKSSQGRLESFEPYAT 345  
 Db 296 DEGLISFLVKGNGNEDRVTKAIEKTSANKSSQGRLESFEPYAT 339

# RESULT 8

E75117  
 DNA repair protein rad2 (rad2) PAB1877 - Pyrococcus abyssi (strain Ozeay)  
 C:Species: Pyrococcus abyssi  
 C:Date: 20-Aug-1999 \*sequence\_revision 20-Aug-1999 \*text\_change 20-Jun-2000

C:Accession: E75117  
 R:anonymous, Genoscope  
 Submitted to the EMBL Data Library, July 1999  
 A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chr  
 A:Reference number: A75001  
 A:Accession: E75117  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-343 <KAW>  
 A:Cross-references: GS:A124828; GS:A105836; NID:95459067; PID:CA649654.1  
 A:Experimental source: strain Ozeay  
 C:Genetics:  
 A:Gene: PAB1877  
 C:Superfamily: conserved hypothetical protein YKL113c

Query Match 32.8%; Score 636; Db 2; Length 343;  
 Best Local Similarity 41.6%; Pred. No. 2.8e-35;  
 Matches 143; Conservative 58; Mismatches 115; Indels 28; Gaps

QY 19 KEQKFEYFGKRIADVAMS.YQPIVAVGTGNETLTNEAGVYTHSGMNTTILFLA 78  
 Db 12 KEILENLYGKRIADVAMS.YQPIVAVGTGNETLTNEAGVYTHSGMNTTILFLA 71  
 QY 79 GIKPYYVDGKPPDMKQOELAKRYSKRDAKLTLEAVEGDKDAIEKLSKRTVYTRQ 138  
 Db 72 GIKPYYVDGKPPDMKQOELAKRYSKRDAKLTLEAVEGDKDAIEKLSKRTVYTRQ 131  
 QY 139 NEDCKRLRLMGVYVVEAPSAEACALCTINKVAVASDDSLTEGAPRCLRLMP 198  
 Db 132 EDCAKLELMGIPVAPSEGAAYMAKGSVASASODVDS.JIFGPRVIRN.TIT 193  
 QY 199 SSKRIPE-----WFEIVAAVLEELTLMQPIIDLCICGCV-CDISK3ICG 244  
 Db 192 GKRLPGKNYVYKPELLE-----EVKELKTRKTELIMAVIGTDPNGIGIGL 247  
 QY 245 QIALKLIRHOGSIESILEN.LNKD--RYQIPEDAPYQEARLCKEPVWLTDFE 301  
 Db 248 KRALHIVRH-----SKDPLAKFOKSDVDLVAIKFEFLNPPV.TENYVAVKRP 295  
 QY 302 DEGLISFLVKGNGNEDRVTKAIEKTSANKSSQGRLESFEPYAT 345  
 Db 296 DEGLISFLVKGNGNEDRVTKAIEKTSANKSSQGRLESFEPYAT 339

# RESULT 9

A71015  
 probable 5' nuclease - Pyrococcus horikoshii  
 C:Species: Pyrococcus horikoshii  
 C:Date: 14-Aug-1998 \*sequence\_revision 14-Aug-1998 \*text\_change 20-Jun-2000  
 C:Accession: A71015  
 A>Title: Complete sequence and gene organization of the genome of a hyper-thermophilic archaeon  
 A:Reference number: A71009; MUID:98344137; PMID:9579194  
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-343 <KAW>  
 A:Cross-references: GS:A000006; NID:q4102613; PID:BA036251.1; PID:q4102613  
 A:Experimental source: strain 013  
 A:Note: This accession replaces an old accession for a sequence replaced  
 C:Genetics:  
 A:Gene: PH145  
 C:Superfamily: conserved hypothetical protein YKL113c

Query Match 32.6%; Score 632.9; Db 2; Length 343;  
 Best Local Similarity 39.2%; Pred. No. 4.8e-35;  
 Matches 142; Conservative 66; Mismatches 119; Indels 35; Gaps

QY 1 MCIGKLTLLADNAKMKQKFEYFGKRIADVAMS.YQPIVAVGTGNETLTNEAGVYTHSGMNTTILFLA 78  
 Db 1 MCVP-----IDDLVIR-KEIDLEMLYGKRIADVAMS.YQPIVAVGTGNETLTNEAGVYTHSGMNTTILFLA 71

[illegible]

RESULT 10  
 C64480  
 DNA repair protein RAD2 homolog - *Methanococcus jannaschii*  
 C:Species: *Methanococcus jannaschii*  
 C:Date: 13-Sep-1996 #sequence,revision 13-Sep-1996 #text\_change 21-Jul-2000  
 C:Accession: C64480  
 R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, A.; Fein, C.L.; Overberg, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; Peterson, J.D.; Sadow, P.M.; Hinkley, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
 Science 273, 1058-1073, 1996  
 A:Authors: Kaine, B.P.; Borodovsky, M.; Kien, H.P.; Fraser, C.M.; Smith, H.O.; Moore, G.  
 A>Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*  
 A:Reference number: A64300; MUID:9637993; PMID:8688087  
 A:Accession: C64480  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Features: 1-326 <ORF>  
 A:Cross-references: GB:U67585; GB:L77117; NID:q1592088; PIDN:AA89456.1; RID:q1592090;  
 C:Genetics:  
 A:Map position: F0R141334-141314  
 C:Superfamily: conserved hypothetical protein YKR113C

```

Query March 28.9% Score 560.5; DB 2; Length 32%;
Best Local Similarity 38.0%; Pred. No. je 30;
Matches 132; Conservative 71; Mismatches 121; Indels 23; Gaps 4;

QY 1 MGKGLTLLDNMPKAMKEQKFESYFGKRIAVDASMSIQYFLIYVGRGTQETLTNNGE 60
   |||  |||  ||  ||| ||| ||| |||
DB 1 MGVG----PDGLPKNI--ISFDIADKQKAIICGMALVQFLISIRIDSSPLNNKGE 53
   |||  |||  |||  |||  |||  |||  |||  |||

QY 61 VTSHLQGNFRIIRLEAGIKPVYVYEDGKPEDMKQKQELAKRYSKRDATKDLTEAVEYGD 120
   |||  |||  |||  |||  |||  |||  |||  |||
DB 54 ITSANGVGYFTIHLLEND;TPWVDEDEPKRLKTKYVRKKEMKAEIKMEKAIKKEE 113
   |||  |||  |||  |||  |||  |||  |||  |||

QY 121 KDALIELXSKRIYVKTROHNEQCKRLIRIMGVYVYVAPSEAPASGALCINIKVAVASND 180
   |||  |||  |||  |||  |||  |||  |||  |||
DB 114 FEEAPAKVAKRSYLTLPKMWENCKYPLSLMGIPYVPEEGEAOASVYMKARKDWAIVASQD 173
   |||  |||  |||  |||  |||  |||  |||  |||

QY 181 MDSLTFGAPREFLRHIMDPSSKKIIPVMEFGVAKVLEELTNDQFIDLCILGCGCV--CSI 234
   |||  |||  |||  |||  |||  |||  |||  |||
DB 174 YDALLYGAPRYVRNL--TTTKMEPL-LELNEVILFDLRISLDDLIDIAIFNGDTYNGCV 230
   |||  |||  |||  |||  |||  |||  |||  |||

QY 240 KGIGQGATALKLIRHGSJESITLENLNKDRQYIPEDMPQGFARRLFEPNVLDIPIETKWT 294
   |||  |||  |||  |||  |||  |||  |||  |||
DB 221 KGIGKRAYETLVRS-GVAKDVLKK-----EVEYDEIDKRIKIFKEPKYV-DNYSLSK 279
   |||  |||  |||  |||  |||  |||  |||  |||

QY 306 AFDEEGILISFLVKONGFNEDEVYTAIEIKIKS-ANKKSSQGRLESEFX 345

```

```

280 LPOKEGI IKFIVDENDFNVDRAKKHVJDKIYNILANKTKQKRTILDAWEK 326

```

RESULT: 11  
C69085

DNA repair protein Rad2 - *Methanobacterium thermoautotrophicum* (strain Delta H)  
 C:Species: *Methanobacterium thermoautotrophicum*  
 C:Date: 05-Dec-1997 #sequence:revision 05-Dec-1997 #text\_change 24-Sep-1999  
 C:Accession: C69385  
 R:Smith, D.R.; DeGutelle-Stamm, L.A.; DeLonghery, G.; Lee, H.; Duponts, C.; P.  
 . Liu, D.; Spadafora, R.; Vicario, R.; Wang, Y.; Mierzejowski, J.; Gibson, F.  
 K.; S.G. Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve  
 J.; Bacteroid. 179, 7135-7155, 1997  
 M:Title: Complete genome sequence of *Methanobacterium thermoautotrophicum* l  
 A:Reference complete: A69000; MIMD:198037514; PMID:9571463  
 A:Accession: C69085  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shc  
 A:Molecule type: DNA  
 A:Residues: 1-328 <MTH>  
 A:Cross-references: GB:AE000922; GB:AE006666; NID:g2622754; PIDN:AA86106.  
 A:Experimental source: strain Delta H  
 C:Genetics  
 A:Gene: MTH163  
 A:Superfamily: conserved hypochetrical protein: YK113-

[illegible]

RESULT 12  
 569282  
 DNA repair protein PAD2 (pad2) homolog - Archaeoglobus fulgidus  
 A:Species: Archaeoglobus fulgidus  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #tex\_change 24-Sep-1997  
 A:Accession: H69282  
 K:Klenk, H.P.; Cayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Kitchum  
 .; Fleischman, R.D.; Overbeek, K.J.; Lee, N.H.; Sutton, G.G.; Gill, S.F.  
 .; Glodek, A.; Zhou, L.; Venter, J.; Gooday, J.D.; Weidman, J.F.; McDonald  
 .; Nature 1997, 364-370, 1997  
 A:Authors: Overbeek, K.J.; Colton, M.D.; Spriggs, T.; Atliach, P.; Kane, B.  
 .; Smith, H.O.; Woese, C.R.; Venter, J.C.  
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reduc  
 .  
 A:Reference numbers: A69520, PMID:9804943, PMID:9389475  
 A:Accession: H69282  
 A:Status: preliminary: nucleic acid sequence not shown; translation not sh  
 .  
 A:Molecule type: DNA  
 A:Residues: 1-336 <RLF>

A:Cross-references: GB:AE001087; GB:AE000782; NID:q2568410; EIDN:AB90367.1; P1D:q2568410  
 C:Superfamily: conserved hypothetical protein YK1130

## Query Match

27.8%; Score 538.5; DB 2; Length 336;  
 Best Local Similarity 36.7%; Pred. No. 9.2e-29;

## Matches

125; Conservative 67; Mismatches 118; Indels 31; Gaps 7;

QY 19 KEQKESYFGKRIADAMSIYOFILYVGRKGTETLNAGEVTSILQGMNRTIFLLIA 78  
 DB 12 EEVELETFSGKRIADADANTLYOFISITRQDGPPLDSDGRIIISLISGLIYRSMNVEV 71  
 QY 79 GIKPYVVEGKPPPMKKKQELAKYKSDRDAKDTLEAVEGDDAKIKLKRIVKTRCH 138  
 DB 72 GIRPVYFVGDPPEKKAKLEERKKRRFAEEMIALQNDKDA-KKYQAQAGRCVSEY 130  
 QY 139 NEDCKRLRLMGVVPVEAPSEAEACALCINDKVPVASEDMDSLTFFGAPRLRLHMDP 198  
 DB 131 VDSAKTLLSYMGIPFVDAPEGEQAAMAKGCVETYSQDYDSLTFGSPRCARNLAI 190  
 QY 199 SSKKIP-----VMEFDVAKYLELELTMOFDLCILGCCDCDS-KGIGSQ 245  
 DB 191 GKRIKPKNNVYVDPKPEITIIIESN-----LKRIGLTRQDLIDATVGTGVEGKRYGVK 246  
 QY 246 TALKLRQHGSIESTLENL--NKORYQIPSDWYQAFARLFEKPNVTLDIPEKMTAPDE 303  
 DB 247 KALNYIKTYGDIFFALKALKVNIH-----VEEIRNFELNPPVTDY-KIFFREEDP 297  
 QY 304 EGIISFLVNDGNFEDRYTKAIEKSKAKNSGGRLESEF 344  
 DB 298 EKAIEFLCEHDFSRERVEKALEKLA--KSTQATLERMP 336

## RESULT 13

H72765

Probable flap endonuclease-1 APE0115 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C:Date: 20-Aug-1999 \*sequence\_revision 20-Aug-1999 \*text\_change 20-Jun-2000

C:Accession: H72765

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Hatahara, Y.; Jin-no, K.; Takahashi, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; K

DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, *Aeropyrum*

A:Reference number: A72450; MIM:99310339; PMID:10382966

A:Accession: H72765

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-401 <KAW>

A:Cross-references: DBJ:AF000058; NID:q5103388; P1DN:BA079026.1; P1D:q5103355

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE0115

C:Superfamily: conserved hypothetical protein YK1130

## Query Match

26.5%; Score 514.5; DB 2; Length 401;  
 Best Local Similarity 34.6%; Pred. No. 4.6e-27;

## Matches

119; Conservative 74; Mismatches 132; Indels 19; Gaps 6;

QY 15 EKAKKQKPESTYFKRIKIVDASMSITQFLIVKRTGMSILTNAGEVTSILQGMNRTIF 74  
 DB 62 PEAREVELRALSGYVALADAYNMLOYPIITAIROPDGPILIDREGHYTSHLSCHETPTN 121  
 QY 75 LLEAGIKPVYVEGKPPPMKKKQELAKYKSDRDAKDTLEAVEGDDAKIKLKRIVKTRCH 134  
 DB 122 IYBEIKTPYVFDGKPPPMKKKQELAKYKSDRDAKDTLEAVEGDDAKIKLKRIVKTRCH 134  
 QY 135 TQNNEDCKRLRLMGVVPVEAPSEAEACALCINDKVPVASEDMDSLTFFGAPRLRLHMDP 194  
 DB 102 TSDVFESEKELIDANGKPMVDAPEGEQAAMAKGCVETYSQDYDSLTFGSPRCARNLAI 245  
 QY 195 LWDPSKAPVYE-----FCVAKYLELELTMOFDLCILGCCDCDS-KGIGSQ 244  
 DB 242 LAITGRRLKPGSDVVEIKETITELDELKSLG-TREQLAVGILLOTCTNFGVGRVGP 301

QY 245 QTALLILQHGSIESTLENLKKDRYQIEPWY-QEARRLFEKPNVTLDIPEKMTAPDE 303  
 DB 302 KALNIVKSLGDPKPKVIVASVPRGY---DQVYLRKVEYHLNPPVTDY-KIFFREEDP 297  
 QY 304 EGIISFLVNDGNFEDRYTKAIEKSKAKNSGGRLESEF 344  
 DB 357 DKVEILVERHDPNFERERALKERLAKYKSKLGRQSLDMP 400

## RESULT 14

F90158

DNA repair endo/exonuclease FEN-1 (RAD2) (tag2) [imported] - Sulfolobus sol

C:Species: Sulfolobus solfataricus

C:Date: 24-May-2001 \*sequence\_revision 24-May-2001 \*text\_change 15-Jun-2001

C:Accession: F90158

R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, V.; Alward, G.; Aways

Jong, T.; Joffiles, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.

arrett, R.A.; Ragan, M.A.; Sengen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A:Description: Sulfolobus solfataricus complete genome.

A:Reference number: A59139

A:Accession: F90158

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-302 <KOR>

A:Cross-references: GB:AE006641; NID:q181333; P1DN:AAK40525.1; GSPDB:GN00

C:Genetics:

A:Gene: rad2

C:Superfamily: conserved hypothetical protein YK1130

## Query Match

26.1%; Score 506; DB 2; Length 302;  
 Best Local Similarity 37.9%; Pred. No. 1.2e-26;

## Matches

118; Conservative 58; Mismatches 97; Indels 38; Gaps

QY 59 GEVTHIQMFRITRILLIACGKPYVYVDPKPPPMKKKQELAKYKSDRDAKDTLEAVE 303  
 DB 5 GRVTHLSGLFRTINILEGVPIVYVDPKPPPMKKKQELAKYKSDRDAKDTLEAVE 303  
 QY 119 GCDKAIHKISKRTVYVYQHNEDEKRLKMGVVPVEAPSEAEACALCINDKVPVAVAS 198  
 DB 65 GRIEELKRYSOAILRLSNINVEESKILRAMSGPIVQAPNSGFAFVYLNKLGSLMAAS 198  
 QY 175 EMDSLTFFGAPRLRLHMDPSKIP-----VMEVDVAKYLELELTMOFDLCILGCCDCDS 244  
 DB 125 QVDAITLFGAKRLVNIITGRKIKNDVYVVEIKPILITEL--LKKLQITREOITDI 190  
 QY 228 GIIICCDY-CDSIKGICGTALIKLRQHGSIESTLENLKKDRYQIEPWY-QEARRLFEKPNV 303  
 DB 183 GIIISTINPDIIRIGTSPERAKIITKKKRIKAKVEYGTISKDINELD----- 303  
 QY 278 QKARKFKPEPNVTLDIPIKMTAHEHGTISFLVKNQFNEDRYTKAIEKSKAKNSGGRLESEF 344  
 DB 233 -ELRGLFNPVYVVEAEADLNENPQND-LNILEYEHNFSEERKNGRIELTKAIKAKG 344  
 QY 334 KESQGLSESEF 344  
 DB 292 ASQGMQIDMP 302

## RESULT 15

F84290

DNA repair protein (imported) - Halobacterium sp. NRC-1

C:Species: Halobacterium sp. NRC-1

C:Date: 02-Feb-2001 \*sequence\_revision 02-Feb-2001 \*text\_change 16-Feb-2001

C:Accession: F84290

R:Katz, M.V.; Kennedy, S.P.; Matsuda, S.G.; Bergquist, B.; Fan, M.; Shukla, H.

Leibhauser, B.; Keller, K.; Cruz, R.; Hansen, M.J.; Hough, D.W.; Maddocks

Jung, K.H.; Alam, M.; Freitas, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Katz, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Fbhardt, H.; LC

A:Title: Genome sequence of Halobacterium species NRC-1

A:Reference number: A81160; MIM:20504483; PMID:11016953

A:Accession: F84290



A: Status: preliminary  
 A: Molecule type: DNA  
 A: Residues: 1-327 <STO>  
 A: Cross-references: CB:AE004437; NID:q10580870; PIDN:AA319630.1; GSPD3:GN00138  
 C: Genetics:  
 A: Gene: rad2  
 C: Superfamily: conserved hypothetical protein YKL113c

Query Match 19.0%; Score 369; D5 2; Length 327;

Best Local Similarity 31.8%; Pred. No. 1.9e-17;

Matches 106; Conservative 50; Mismatches 139; Indels 28; Gaps 9;

```

QY 17 AKKEQKESYFGRIADVASMTYQFLIYVGRTGNETLTNEAGVTSHTQGFNRTIRL 75
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 11 ALEETPFADLEGSVAADAHNMILKYLTTVQWAGADVTTSDGTEVANLWCAVGGIPK 70
QY 76 LEAGIKPYVFPDGKPPDKKQFLAKRYSKRDDATKDLTFAVEVGDKDALFKLSKRTVXYT 135
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 71 FEHGILTPVFWMDGGVTEIKDDEIADREQREYEEQLDAREAGDAEAARLDAFTQRLT 130
QY 136 RQHNEDCKRLRLKMGVPYVEAPSEAEACALCINUKVFAVASSEMDSLTFCAGTFLR 194
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 131 PTIETTRLEFDLDIPQVEAPAEGEAAVMTRIDPAVDYAGSDDYDCLLSPTLRQ 190
QY 195 LMDPSKRIPVMEFPAKAVLEFELEITMIOFLCICLOGCDYCDISKIGGQFALKLRQH 254
   || | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 191 L--TSSGHEIMFDA--TCAEHDLTWEQLVDVGLCGTDFNPGIDGFGPTALDAGCH 246
QY 255 GSLESTL---ENL-NKDRYQIPEDWPIQFARRLFKSPNVTLDI---PELKNATATDEEG 306
   || | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 247 GDLMVLAEGEHHVAGDR-----LRELFNPDVTDVYIDFV---SFAIDAA 292
QY 307 ISFLVNDGFNEDRYTKAIEKIKSAKNSQGR 339
   || | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 293 RAVVTDEMEVDADAVARGFERIDMAAAGTGLNR 325

```

Search completed: October 22, 2003, 12:43:25  
 Job Time : 44 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 22, 2003, 12:20:40 ; Search time 23 Seconds

(without alignments)  
774,918 Million cell updates/sec

Title: US-09-805-311-2  
Perfect score: 1939  
Sequence: 1 MGICLTKLADNAPKAKE.....SDKSKAAANKKAKGCKKK 379

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues  
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 3%  
Maximum Match 10%  
Listing first 45 summaries

Database: SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1032	53.2	380	1	FEN1_HUMAN
2	1023.5	52.8	380	1	RAD2_SCHPO
3	1014	52.3	378	1	FEN1_MOUSE
4	938	48.4	382	1	RAD2_YEAST
5	648.5	33.4	340	1	FEN1_PRRU
6	636	32.8	343	1	FEN1_PRRAB
7	632.5	32.6	343	1	FEN1_PRRHO
8	628	32.4	346	1	FEN1_PRRAB
9	622	32.1	348	1	FEN1_PRRAB
10	586.5	30.2	351	1	FEN1_METAC
11	582	30.0	351	1	FEN1_SULTO
12	562	29.0	345	1	FEN1_SULTO
13	560.5	28.9	326	1	FEN1_MENJA
14	560.5	28.9	326	1	FEN1_MENJA
15	542	28.0	328	1	FEN1_MENJA
16	548.5	27.8	336	1	FEN1_MENJA
17	514.5	26.5	351	1	FEN1_AERPE
18	468	24.1	336	1	FEN1_THERC
19	446	23.0	335	1	FEN1_THERC
20	369	19.0	327	1	FEN1_HAINI
21	260.5	13.4	571	1	EXO1_SCHPO
22	245.5	12.7	496	1	RAD2_SCHPO
23	232.5	12.0	1031	1	XPG_HUMAN
24	220.5	11.4	1186	1	XPG_MOUSE
25	220	11.3	1170	1	XPG_MOUSE
26	218	11.2	702	1	EXO1_YEAST
27	204.5	10.5	430	1	RAD2_SCHPO
28	204.5	10.5	1112	1	RAD2_SCHPO
29	200	10.3	1196	1	XPG_XENLA
30	184.5	9.5	834	1	DPO1_THERC
31	176.5	9.1	586	1	DPO1_SYNY3
32	173	8.9	832	1	DPO1_THERC
33	172.5	8.9	834	1	DPO1_THERC

## ALIGNMENTS

RESULT 1	ID	FEN1_HUMAN	STANDARD:	PRG:	380 AA
AC	P39748	FEN1_HUMAN			
DT	01-FEB-1995 (Rel. 31, Created)				
DT	01-FEB-1995 (Rel. 31, Last sequence update)				
DT	15-SEP-2003 (Rel. 42, Last annotation update)				
DE	Flap endonuclease-1 (EC 3.1.1.1) (Maturation factor 1) (MF1).				
OS	FEN1.				
OS	Homo sapiens (Human).				
OS	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OS	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.				
OS	NCBI_TaxID=9606;				
SN	11				
SN	SEQUENCE FROM N.A.				
SN	MDL:194277893; PubMed=805745;				
SN	Murray J.M., Tavassoli M., Al-Hadidy R., Sheldrick K.S.,				
SN	Lehman A.R., Carr A.M., Watts P.Z.,				
SN	"Structural and functional conservation of the human homolog of the				
SN	Schizosaccharomyces pombe rad2 gene, which is required for chromosome				
SN	segregation and recovery from DNA damage."				
SN	Mol. Cell Biol. 14:4876-4888(1994).				
SN	12				
SN	SEQUENCE FROM N.A.				
SN	TISSUE-Lymphocytes;				
SN	Lieber M.R., Hsieh C.L., Harrington J.J.,				
SN	Submitted (FEB-1995) to the EMBL/Genbank/DBJ databases.				
SN	13				
SN	SEQUENCE FROM N.A.				
SN	Rieder M.J., Livingston R.J., Bran A.C., Montoya M.A., Chung M.-M.,				
SN	Miyamoto K.E., Nguyen C.P., Pohl C.L., Robertson P.D.,				
SN	Schackwitz W.S., Sherwood J.K., Wittek L.A., Nickerson D.A.,				
SN	Submitted (JUN-2002) to the EMBL/Genbank/DBJ databases.				
SN	14				
SN	SEQUENCE FROM N.A.				
SN	Lamertine J.F., McGready P.M., Coleman M., Stowinski F., Adamson A.W.,				
SN	Burkhardt-Schultz K., Gordon L., Kyle A., Ramirez M., Sellwage S.,				
SN	Phan H., Velasco N., De L., Kestler N., Terry A., Ganes J.,				
SN	Dadgar D., Poundstone P., Christensen M., Georgescu A., Avila J.,				
SN	Lin S., Altix C., Andreise T., Traubheim M., Amico-Keller G.,				
SN	Coefield J., Duarte S., Lucas S., Bruce K., Thomas P., Quan G.,				
SN	Krommiller B., Arellano A., Montgomery M., Qu D., Nolan M., Trong S.,				
SN	Kobayashi A., Olsen A.S.,				
SN	Carraro A.V.,				
SN	Submitted (MAY-1998) to the EMBL/Genbank/DBJ databases.				
SN	15				
SN	SEQUENCE FROM N.A.				
SN	TISSUE-Lung;				
SN	MDL:194277893; PubMed=1477932;				
SN	Strasser R.L., Pelting E.A., Grouse L.H., Derge C.C.,				
SN	Klausner R.J., Collins F.S., Wagner L., Shenmen G.M., Schuler G.D.,				
SN	Altshul S.F., Zeeberg B., Buelow K.B., Schaefer C.F., Bhat N.K.,				
SN	Hopkins R.F., Jordan H., Moore T., Max S., Wang J., Hsieh F.,				
SN	Diachenko L., Marusina K., Farrer A.A., Rubin G.M., Hong L.,				
SN	Stapleton M., Shares M.B., Rosalito M.F., Casavant T.L., Scheetz T.E.,				
SN	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,				

RA Raha S.S., Joquelliano N.A., Peters G.J., Abramson R.D., Muliady S.J.,  
 RA Rosak S.A., McWen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.R.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska J., Smalins D.E.,  
 RA Schercher A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN 16)  
 RP PARTIAL SEQUENCE.  
 RX MEDLINE-95050647; PubMed-7961795;  
 RA Robins P., Pappin D.J., Wood R.D., Lindahl T.;  
 RT "Structural and functional homology between mammalian DNase IV and  
 RT the 5'-nuclease domain of *Escherichia coli* DNA polymerase I.";  
 RL J. Biol. Chem. 269:28535-28539(1994).  
 RN 17)  
 RP ACETYLATION.  
 RA Hasen S., Stucki M., Hassa P.O., Imhof R., Gehrig P., Hunziker P.,  
 RA Hubscher J., Hottiger M.O.;  
 RT "Regulation of human flap endonuclease-1 activity by acetylation  
 RT through the transcriptional coactivator p300.";  
 RL Mol. Biol. Cell 7:1221-1231(2001).  
 CC -1- FUNCTION: Endonuclease that cleave the 5'overhanging flap DNA  
 CC structure that is generated by displacement synthesis when DNA  
 CC polymerase encounters the 5'end of a downstream Okazaki fragment.  
 CC It fails to cleave other DNA structures, including 3'flaps and  
 CC single stranded DNA (by similarity).  
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -1- PTM: Acetylated by p300.  
 CC -1- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. FEN1  
 CC SUBFAMILY  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC EMBL: X76771; CAAS4166.1;  
 DR EMBL: L37374; AAA91331.1;  
 DR EMBL: AF523117; AAM74238.1;  
 DR EMBL: AC004770; AAC23394.1;  
 DR EMBL: BC000323; AAH00323.1;  
 DR PIR: A56531; A56531.  
 DR HSSP: Q58839; IAT6.  
 DR GeneW: HGNC:3650; FEN1.  
 DR GK: P39748;  
 DR KIM: 600393;  
 DR GO: GO:0003686; F:DNA repair enzyme; TAS.  
 DR GO: GO:0003690; F:double-stranded DNA binding activity; TAS.  
 DR GO: GO:0008309; F:double-stranded DNA specific endonuclease activity; TAS.  
 DR GO: GO:0004519; F:endonuclease activity; TAS.  
 DR GO: GO:0004527; F:endonuclease activity; TAS.  
 DR GO: GO:0006260; P:DNA replication; TAS.  
 DR GO: GO:0006302; P:double-strand break repair; TAS.  
 DR GO: GO:0009653; P:UV protection; TAS.  
 DR InterPro: IPR002421; 5\_3\_exonuclease.  
 DR InterPro: IPR00513; Exo\_N\_1.  
 DR InterPro: IPR003584; HNH\_2.  
 DR InterPro: IPR006086; XP3\_1.  
 DR InterPro: IPR006085; XPG\_N.  
 DR InterPro: IPR006084; XPGC\_Pad.  
 DR Pfam: PF00867; XPG\_1.  
 DR Pfam: PF00752; XPG\_N.  
 DR PRINTS: PR00553; XPGRADSUPR.  
 DR SMART: SK00279; HNH2; 1.

DR SMART: SK00484; XPG1; 1.  
 DR SMART: SK00485; XPGN; 1.  
 DR PROSITE: PS00841; XPG\_1; 1.  
 DR PROSITE: PS00842; XPG\_2; 1.  
 KW Hydrolase; Nuclease; Endonuclease; Nuclear protein; Acetylation.  
 FT DOMAIN 1 104 N-DOMAIN.  
 FT DOMAIN 122 253 T-DOMAIN.  
 SC SEQUENCE 380 AA; 42593 MW; 5145F2E0E5759205 CRC64;  
 Query Match 53.28; Score 1032; DB 1; Length 380;  
 Best Local Similarity 53.98; Pred. No. 1,86-61;  
 Matches 206; Conservative 63; Mismatches 98; Indels 19; Gaps  
 QY : MGKIKITKLADNAPKARKKQKFESEYGRKCAVDASKSYVQGVGRTQKFTVEACE  
 Db : MGKIGIAKLADAPNAIRNDIKSTFGKRVALLDANSITQFLIAY-RGSDPIQNEEG  
 QY : VTSHLQMPKRTLRLEAGIPVYVQGRKPKKQKQCAKRYSKQDARTDCTEAVEVGG  
 Db : VTSHLMQMPKRTLRKMKNGIKPVYVQGRKPKKQKQCAKRYSKQDARTDCTEAVEVGG  
 QY : 121 KVALERKSKIVVVTROHNEDEKRLRLKRVPEVAEASEAFACGALCTDCKPFAVASED  
 Db : 120 KQVEKRTKRLVVTROHNEDEKRLSLMGIPVLCASEAEASTALYKAGKYAAATID  
 QY : 181 MSLTEGAPRFTCHLMQPSKRTFVMEFDYAKVLESELTMDFIDLCIGGCDYDSIK  
 Db : 180 MDCITIGSPVLMKHLTASEAKKLPIDFHLRSLQEGNQPFVDCILIGSDYIESIR  
 QY : 241 GIGGQIAKLIROHGESILENKKNDKQVTPEDKRYQFARLFRKPNNTL--ITELKW  
 Db : 240 GIGPKRAVDILQKHSILEYVRLDPRKYPVPMWLMKEKHQFLFEV--LDSEVELKW  
 QY : 259 TAPDEGLISFLKDNKFNEDPRTVKIEKSKSKNKSQGLSPFPPTAFAPLKRKE  
 Db : 259 SEPRERELIKPKQKQKQFSEKRSRSGYKRSKRSKQSGIQGLDPPKVTGSLSS-AKRKE  
 QY : 359 TSDK--TSKAANKKTIACQKXK 379  
 Db : 358 PEPKGRSKKKA---KTFACQKFK 377  
 RESULT 2  
 RAD2\_SCHPO STANDARD: PRT; 350 AA.  
 ID RAD2\_SCHPO  
 AC P19750;  
 DT 01-FEB-1995 (Ref. 31, Created)  
 DT 01-FEB-1995 (Ref. 31, Last sequence update)  
 DT 26-FEB-2003 (Ref. 41, Last annotation update)  
 DE DNA repair protein Rad2.  
 GN RAD2 OR SPAC366.06C.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 CC Schizosaccharomycetes.  
 CX NCBI\_TaxID 4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-94277093; PubMed-8079885;  
 RA Murray J.M., Tavassoli M., A. Hartley R., Sheidrick K.S.,  
 RA Lehmann A.R., Carr A.M., Watts P.Z.;  
 RT "Structural and functional conservation of the human homolog of the  
 RT Schizosaccharomyces pombe rad2 gene, which is required for chromosome  
 RT segregation and recovery from DNA damage.";  
 RL Mol. Cell. Biol. 14:4878-4888(1994).  
 RN 12)  
 RP SEQUENCE FROM N.A.  
 RX STRAIN-972;  
 RX MEDLINE-21846401; PubMed-11859360;  
 RA Wood V., Gilliam P., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Santos J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown J., Brown S., Chillingworth T., Churche C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Delwail T., Fraser A.,

RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgeson G.,  
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
RA Mooney P., Moyle S., Mungall K., Murphy L., Niblett D., Odell C.,  
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
RA Woodward J., Volckaert G., Aert R., Robben J., Grymopriaz B.,  
RA Weljens I., Vastreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,  
RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,  
RA Zger P., Zimmermann W., Weiler H., Wembut R., Purnelle B.,  
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaire V., Mottier S.,  
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Rhode G.,  
RA Daga R.R., Cruzado I., Jimenez J., Sanchez M., del Rey F., Benito J.,  
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
RA Cerutti L., Jove T., McCombie W.R., Paulsen J., Potashkin A.,  
RA Shpakovski G.V., Ussery D., Bartell B.G., Nurse P.,  
RT \*The genome sequence of Schizosaccharomyces pombe\*;  
RL Nature 415:87-880(2002).  
CC -!- FUNCTION: REQUIRED FOR THE FIDELITY OF CHROMOSOME SEPARATION AT  
CC MITOSIS AND WHICH IS ALSO INVOLVED IN THE RESPONSE TO DNA DAMAGE.  
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).  
CC -!- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. FEN1  
CC SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/annouco/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: X77041; GA836992.1; -  
DR EMBL: Z99167; CAB15282.1; -  
DR PIR: A56054; A56054.  
DR HSPD: Q58839; 1A76.  
DR GeneDB: Spombe; SPAC3G6.06c; -  
DR InterPro: IPR003513; EXO\_N.1.  
DR InterPro: IPR003584; HNH\_2.  
DR InterPro: IPR006086; XPG\_2.  
DR InterPro: IPR005085; XPG\_N.  
DR InterPro: IPR006084; XPGC\_Rad.  
DR Pfam: PF00867; XPG\_1; 1.  
DR Pfam: PF00752; XPG\_N; 1.  
DR PRINTS: PR00853; XPGRADSUPER.  
DR SMART: SM00279; HNH2; 1.  
DR SMART: SM00484; XPG1; 1.  
DR SMART: SM00485; XPGN; 1.  
DR PROSITE: PS00841; XPG\_1; 1.  
DR PROSITE: PS00842; XPG\_2; 1.  
KW DNA repair; Hydrolase; Nuclease; Endonuclease; Nuclear protein.  
FT DOMAIN 1 105 N-DOMAIN.  
FT DOMAIN 223 254 I-DOMAIN.  
SQ SEQUENCE 380 AA; 42866 MW; 3604065D93934CBF CRC64;

DR 121 AENVDSFANRYKVTROHNDENKRLLEIMGIPFNAPCEAEKOCALASGVYVAASED  
QY 181 MDSLTGGAPFLPHLMDPSSKIPVAFVYAKLELELTMDPIDIILCCGCDYDSIK  
DB 181 MCLTCAVAVLRLHLPFRQKRPISBYINERALKNLDSVDFVOLDLGGDCEPR  
QY 241 GIGGCTALKLIRHGSISTELENLNDRQYIPEDWFOEARLFEPPNTLDP-----EL  
DB 241 GVPFARAVELIRGYTLDFKFKADSKSPYIPEDWYEDARLFLDLAEV---LPGSEIEL  
QY 297 KWIAPPEFGISLIVNDGNENRVRKALEKISAKNKSOGRLSEFFKPTATSAKLK  
DB 298 KWSPPDGLIGFLVEKRGNEBVRVLGIRLEKASKTIPGRDLSFKP--VPSSPKAP  
QY 357 KETSDYTSKAAANKKTKAGCKK 379  
DB 356 VDKSKOSAKRKDKSKKGGSKK 378  
RESULT 3  
FEN1\_MOUSE STANDARD; FRT; 478 AA.  
AC P39749;  
DT 01-FEB-1995 (Rel. 31, last sequence update)  
DT 26-FEB-2003 (Rel. 41, last annotation update)  
DE Flap endonuclease-1 (EC 3.1.1.7).  
OS FEN1 OR FEN-1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Clariata; Vertebrata; Eucelostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurinae; Muridae; Murinae; Mus.  
OC NCBI\_TaxID:10090;  
PN 1;  
RP SEQUENCE FROM N.A. AND SPOCKEN OF 260 275 AND 731-352.  
RC STEADY-BALANCE;  
RX MEDLINE:95015546; PubMed:7927739;  
KA Harrington J.J., Cleber K.R.,  
RT "Functional domains within FEN-1 and RAD2 define a family of  
RT structure-specific endonucleases: implications for nucleotide excision  
RT repair";  
RL Genes Dev. 8:1344-1355(1994).  
CC -!- FUNCTION: Endonuclease that cleave the 5'overhanging flap  
CC structure that is generated by displacement synthesis when DNA  
CC polymerase encounters the 5'end of a downstream Okazaki fragment.  
CC It fails to cleave other DNA structures, including 3'flaps and  
CC single stranded DNA.  
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).  
CC -!- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. FEN1  
CC SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collabora  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commerc  
CC entities requires a license agreement (See <http://www.isb-sib.ch/annou>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: L26320; AAC37664.1; -  
DR PIR: A53730; A53730.  
DR HSPD: Q58839; 1A76.  
DR MGD: MG1102779; FEN1.  
DR InterPro: IPR002421; 5\_3\_exonuclease.  
DR InterPro: IPR005113; EXO\_N.1.  
DR InterPro: IPR005584; HNH\_2.  
DR InterPro: IPR006096; XPG\_1.  
DR InterPro: IPR006085; XPG\_N.  
DR InterPro: IPR006084; XPGC\_Rad.  
DR Pfam: PF00867; XPG\_1; 1.  
DR Pfam: PF00752; XPG\_N; 1.  
DR PRINTS: PR00853; XPGRADSUPER.  
DR SMART: SM00279; HNH2; 1.  
DR SMART: SM00484; XPG1; 1.

SMART: SM00485; XPGN: 1.  
 DR PROSITE: PS00841; XPG\_1: 1.  
 DR PROSITE: PS00842; XPG\_2: 1.  
 KW Hydrolyase, Nuclease, Endonuclease; Nuclear protein.  
 FT DOMAIN 1 102 N-DOMAIN.  
 FT DOMAIN 120 251 1-DOMAIN.  
 SQ SEQUENCE 376 AA: 42314 MW: 827946 HDBPCEP39 CR064;

Query Match 52.38; Score 10.4; DB 1; Length 376;  
 Best Local Similarity 53.28; Pred. No. 2,9e-60;  
 Matches 202; Conservative 69; Mismatches 101; Indels 8; Gaps 6;

QY 1 MGIGKTLKLDANAPKAKKQKFESEYFGKRIADVASMSTIYQFLIVGRGKMETLNEAVE 60  
 D 1 MGIGKTLKLDANAPKAKKQKFESEYFGKRIADVASMSTIYQFLIVGRGKMETLNEAVE 59  
 QY 61 VTSHLQGMENRTITLLENGIKPYVYEDGKPPDKKQKQELAKRYSKRDQATKDLTEAVEVSD 120  
 D 60 TIS-LMGMYFRITR-MENCIKPYVYEDGKPPDKKQKQELAKRYSKRDQATKDLTEAVEVSD 117  
 QY 121 KDAIEKLSKRTVKTROHNDCKRLRLMGVPVVEAFSEAEACALCIDKVFVAVSEF 180  
 D 118 EEEVEKTKRLVKTROHNDCKRLRLMGVPVVEAFSEAEACALCIDKVFVAVSEF 177  
 QY 181 MDLITFGAPRFLRLMDPSSKKIPWEEDVAKYLEELTMDQFDICLQGDYCSIR 240  
 D 178 MDCLTFGSPVLMRLTLSEAKKLPIDFHLISVLOELGMDQFVDCILQGDYCSIR 237  
 QY 241 GIGGQATALKLIRHGSISLELNLRDYOIPEDMPYQEARLFKEPNVT-LDIPELKWI 299  
 D 238 GIGGQATALKLIRHGSISLELNLRDYOIPEDMPYQEARLFKEPNVT-LDIPELKWI 297  
 QY 300 APDEGLISFLYKDNQGFEDRYTKALEIKSNKSSQGRLESEFFKPTATSPCKRKET 359  
 D 298 EPNEELIYKFMCGKQSEERISGVKRLSKRSQGSTQGRLDQFVYTGSLSS-AMKPEP 356  
 QY 360 SDKTSKAAANKKTKACGKK 379  
 D 357 E---PKGPARKKATGSAK 373

RESULT 4  
 RA27\_YEAST  
 ID RA27\_YEAST STANDARD; PRT: 382 AA.

AC P26793; 1  
 DT 01-NOV-1992 (Rel. 23, Created)  
 DT 01-APR-1992 (Rel. 23, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Structure specific endonuclease RAD27.  
 GN RAD27 OR RTN1 OR YKL113C OR YKJ510.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 CX NCB1\_TaxID:4932;  
 FN 11  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE:32221699; PubMed:1561835;  
 RA Jaquier A., Legrain P., Dujon B.;  
 RT Sequence of a 10.7 kb segment of yeast chromosome XI identifies the  
 RT APN1 and the BAP1 loci and reveals one tRNA gene and several new open  
 RT reading frames including homologs to RAD2 and kinases.  
 RL Yeast 5:121-132(1992).  
 RP CHARACTERIZATION.  
 RC STRAIN:3A1B/C;  
 RX MEDLINE:95011546; PubMed:7926735;  
 RA Harrington J.J., Lieber M.R.;  
 RT "Functional domains within FEN-1 and RAD2 define a family of  
 RT structure-specific endonucleases: implications for nucleotide  
 RT excision repair."  
 RL Genes Dev. 8:1344-1355(1994).  
 RP CHARACTERIZATION.

RX MEDLINE:95113773; PubMed:7814325;  
 RA Reagan M.S., Pittenger C., Siede W., Friedberg E.C.;  
 RT "Characterization of a mutant strain of Saccharomyces cerevisiae with  
 RT a deletion of the RAD27 gene, a structural homolog of the RAD2  
 RT nucleotide excision repair gene."  
 RL J. Bacteriol. 177:364-371(1995).  
 CC 1- FUNCTION: SINGLE-STRANDED DNA ENDONUCLEASE (PROBABLE).  
 CC 1- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC 1- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. FEN1  
 SUBFAMILY.

CC This SMARTS-PAGE entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
 CC the European Bioinformatics Institute. There are no restrictions on  
 CC use by non-profit institutions as long as its content is in no  
 CC modified and this statement is not removed. Usage by and for commer  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/about>  
 CC or send an email to [license@sib.ch](mailto:license@sib.ch)).

DR EXHL: S93804; A831998.1;  
 DR EMBL: Z28113; CA81953.1;  
 DR PIR: S22267; S22267.  
 DR HSP: Q58399; IAV6.  
 DR GK: P26793;  
 DR SGD: S0001596; RAD27.  
 DR GO: GO:0005634; Cnucleus; IMP.  
 DR GO: GO:001708; Fflap endonuclease activity; IDA.  
 DR GO: GO:000731; P:DNA repair synthesis; IMP.  
 DR InterPro: IPR002421; 5.3\_exonuclease.  
 DR InterPro: IPR000513; EXO\_N\_L.  
 DR InterPro: IPR003584; HHM\_2.  
 DR InterPro: IPR006086; XPG\_1.  
 DR InterPro: IPR006085; XPG\_N.  
 DR InterPro: IPR006084; XPG\_Nad.  
 DR Pfam: PF00867; XPG\_1; 1.  
 DR Pfam: PF00752; XPG\_N; 1.  
 DR PRINTS: PR00853; XPGRADSCPER.  
 DR SMART: SM00279; HHz2; 1.  
 DR SMART: SM00484; XPG1; 1.  
 DR SMART: SM00485; XPGN; 1.  
 DR PROSITE: PS00841; XPG\_1; 1.  
 DR PROSITE: PS00842; XPG\_2; 1.  
 DR DNA repair: Hydrolyase, Nuclease; Endonuclease; Nuclear protein.  
 FT DOMAIN 1 120 251 N-DOMAIN.  
 FT DOMAIN 1 105 N-DOMAIN.  
 SQ SEQUENCE 382 AA: 43279 MW: 1554808720 2109C CR064;

Query Match 48.4%; Score 9.9; DB 1; Length 382;  
 Best Local Similarity 49.5%; Pred. No. 3.2e-55;  
 Matches 190; Conservative 69; Mismatches 129; Indels 16; Gaps

QY 1 MGIGKTLKLDANAPKAKKQKFESEYFGKRIADVASMSTIYQFLIVGRGKMETLNEAVE 60  
 D 1 MGIGKTLKLDANAPKAKKQKFESEYFGKRIADVASMSTIYQFLIVGRGKMETLNEAVE 59  
 QY 61 VTSHLQGMENRTITLLENGIKPYVYEDGKPPDKKQKQELAKRYSKRDQATKDLTEAVEVSD 120  
 D 60 TIS-LMGMYFRITR-MENCIKPYVYEDGKPPDKKQKQELAKRYSKRDQATKDLTEAVEVSD 117  
 QY 121 KDAIEKLSKRTVKTROHNDCKRLRLMGVPVVEAFSEAEACALCIDKVFVAVSEF 180  
 D 118 EEEVEKTKRLVKTROHNDCKRLRLMGVPVVEAFSEAEACALCIDKVFVAVSEF 177  
 QY 181 MDLITFGAPRFLRLMDPSSKKIPWEEDVAKYLEELTMDQFDICLQGDYCSIR 240  
 D 178 MDCLTFGSPVLMRLTLSEAKKLPIDFHLISVLOELGMDQFVDCILQGDYCSIR 237  
 QY 241 GIGGQATALKLIRHGSISLELNLRDYOIPEDMPYQEARLFKEPNVT-LDIPELKWI 299  
 D 238 GIGGQATALKLIRHGSISLELNLRDYOIPEDMPYQEARLFKEPNVT-LDIPELKWI 297  
 QY 300 APDEGLISFLYKDNQGFEDRYTKALEIKSNKSSQGRLESEFFKPTATSPCKRKET 359  
 D 298 EPNEELIYKFMCGKQSEERISGVKRLSKRSQGSTQGRLDQFVYTGSLSS-AMKPEP 356  
 QY 360 SDKTSKAAANKKTKACGKK 379  
 D 357 E---PKGPARKKATGSAK 373

```

DB      297 INKMSPEKEKELIYLCDDKKFSEERYKSGISRLKGLKSGIQSLDGFYVVPKTEQ 356
QY      351 -SAPLRKETSOKTSGKAANKKTK 373
      357 LAAAKRAOEKRLKLNK-MKNKVTK 379

RESULT: 5
FEN_PYRFU STANDARD: PRT: 340 AA.
AC 093634;
DI 28-FEB-2003 (Rel. 41, Created)
DI 28-FEB-2003 (Rel. 41, Last sequence update)
DI 28-FEB-2003 (Rel. 41, Last annotation update)
DE Flap structure-specific endonuclease (EC 3.1.1.1).
GN FEN OR FEN-1 OR PFI414.
OS Pyrococcus furiosus.
OC Archaea: Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261;
RN 1)
RP SEQUENCE FROM N.A.
RC STRAIN-VCI / DSM 3638 / ATCC 43587 / JCM 8422;
RA MEDLINE:9941585; PubMed:10486095;
RA Dirugiero J., Brown J.R., Beger A.P., Robb E.T.;
RC "DNA repair systems in archaea: mementos from the last universal
RT common ancestor?";
RL J. Mol. Evol. 49:474-484(1999).
RN 12)
RP SEQUENCE FROM N.A.
RC STRAIN-VCI / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weiss R.B., Dunn D.M., Robb E.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome.";
RL Submitted (Feb-2002) to the EMBL/GenBank/DBJ databases.
RN 3)
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RA MEDLINE:98449703; PubMed:9778254;
RA Hostald D.J., Mol C.D., Shen B., Talner J.A.;
RT "Structure of the DNA repair and replication endonuclease and
RL exonuclease FEN-1: coupling DNA and PCNA binding to FEN-1 activity.";
RL Cell 95:135-146(1998).
CC -1- FUNCTION: Endonuclease that cleave the 5'overhanging flap
CC structure that is generated by displacement synthesis when DNA
CC polymerase encounters the 5' end of a downstream Okazaki fragment.
CC Has 5'endo-/exonuclease and 5'pseudo-Y-endonuclease activities.
CC Cleaves the junction between single and double-stranded regions of
CC flap DNA.
CC -1- COFACTOR: Binds 2 magnesium ions per subunit.
CC -1- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. FEN1
CC SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AF013497; AAD01514.1;
DR EMBL: AF010244; AA:81538.1;
DR PIR: T46893; T46893.
DR PDB: 1BA3; 12-JAN-00.
DR HAMAP: MF_00614; 1.
DR InterPro: IPR000513; Exo_N_1.
DR InterPro: IPR003584; HHH_2.
DR InterPro: IPR006086; XPG_1.
DR InterPro: IPR006085; XPG_N.
DR InterPro: IPR006084; XPGC_Rad.
DR Pfam: PF00867; XPG_1; 1.
DR Pfam: PF00752; XPG_N; 1.
DR PRINTS: PR00853; XPGRADSPER.
DR SMART: SM00779; HHH2; 1.

```

```

DR SMART: SM00484; XPG1; 1.
DR SMART: SM00485; XPGN; 1.
DR PROSITE: PS00841; XPG_1; FALSE NEG.
KM Hydrolyase; Nuclease; Endonuclease; Magnesium; Metal-binding;
KT 3D-structure; Complete proteome.
FT METAL 154 154 MAGNESIUM 1 (BY SIMILARITY).
SQ SEQUENCE 340 AA; 34738 MW; D8D3D0F995521F CRC64;

Query Match 33.4%; Score 648.5; DB:1; Length 340;
Best Local Similarity 41.3%; Pred. No. 4,3e-36;
Matches 142; Conservative 62; Mismatches 107; Indels 33; Gaps

QY 19 KEKFESEYFGRKAVDASMSIYQFLIVGRTQMETLTNGEVSITSLQGNFRTIRLLEA
DB 11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 12 KEICLENLYGKRIADALNATYQSLTIRQKQDPLDKSGKRTISHSGFLPRTINMEA
QY 79 GIKFVYVDGKPPQMKQELAKRSKKDKATKLTAVAVQDKNA:PKISKRTYKV:ROH
DB 11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 72 GIKFVYVDGPPPTKKKELEKREAVEEVEKKRREALEKGELEARKYQORATRVNKL
QY 139 NEQKRLIRZMGVAVVAVSFAEACALCINKEVFAVASEDMQSTFGAPFLRLMDP
DB 11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 132 IEQKRLLELMGTVIYASISGMAVAMAKGVYASASQYNSLFGAPRLVRLNLTIT
QY 199 SSKRTF.....-VMEFPAVAFIELTMDQFDTICQGDY-CGSIKQSG
DB 11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 192 GKRLKPKGNVVEIKPELLILE---HYKELKLFKEKLELAVLWGVNPGIKIGTL
QY 245 QTAALKLRQHSISTELENKNC---KVCITHEWVYGLARLPREPVTITDIPELKVTAV
DB 11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 248 KKALETVER-----SKFLAVKQVQSDVITAIKEFFLNIPVT-DNIVLWRDP
QY 302 DEEHLISFVYKQNFNEVTKATKIKSKNKSQGRLEFFK 345
DB 11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 296 DGGTILKTCGFNFSPFRKVNKLRLKKA:KSKQSTLSMK 339

RESULT: 6
FEN_PYRFU STANDARD: PRT: 340 AA.
AC 093634;
DI 28-FEB-2003 (Rel. 41, Created)
DI 28-FEB-2003 (Rel. 41, Last sequence update)
DI 15-SEP-2003 (Rel. 42, Last annotation update)
DE Flap structure-specific endonuclease (EC 3.1.1.1).
GN FEN OR PYR407400 OR PAB1877.
OS Pyrococcus abyssi.
OC Archaea: Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=29252;
RN 1)
RP SEQUENCE FROM N.A.
RC STRAIN-GES / URSAY;
RA PubMed:12622608;
RA Cohen G.N., Barbe V., Flament D., Galperin M., Hellig R., Lecompte O.,
RA Poch O., Prieur D., Querellou J., Ripp R., Thierry J.-C.,
RA Van der Oost J., Weissenbach J., Zivanovic Y., Forterre P.;
RT "An integrated analysis of the genome of the hyperthermophilic
RL archaeon Pyrococcus abyssi.";
RL Mol. Microbiol. 47:1495-1512(2003).
CC -1- FUNCTION: Endonuclease that cleave the 5'overhanging flap
CC structure that is generated by displacement synthesis when DNA
CC polymerase encounters the 5'end of a downstream Okazaki fragment.
CC Has 5'endo-/exonuclease and 5'pseudo-Y-endonuclease activities.
CC Cleaves the junction between single and double-stranded regions o
CC flap DNA (By similarity).
CC -1- COFACTOR: Binds 2 magnesium ions per subunit (By similarity).
CC -1- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASH FAMILY. FEN1
CC SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on

```

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).)

DR EMBL: AJ248285; CAB49654.1; ..  
 DR PIR: E75117; E75117.  
 DR HSSP: Q58839; 1A76.  
 DR HAMAP: MF\_00614; .. 1  
 DR InterPro: IPR000513; Exo\_N.1.  
 DR InterPro: IPR031584; HH2\_2.  
 DR InterPro: IPR006086; XPG\_1.  
 DR InterPro: IPR006085; XPG\_N.  
 DR InterPro: IPR006084; XPG\_Rad.  
 DR Pfam: PF00867; XPG\_1; 1.  
 DR Pfam: PF00752; XPG\_N; 1.  
 DR PRINTS: PR00853; XPGRADSUPER.  
 DR SMART: SM00279; HH2; 1.  
 DR SMART: SM00484; XPG1; 1.  
 DR SMART: SM00485; XPGN; 1.  
 DR PROSITE: PS00841; XPG\_1; FALSE NEG.  
 KM Hydrolyase; Nuclease; Endonuclease; Magnesium; Metal-binding;  
 KW Complete proteome.  
 FI METAL 154  
 SI SEQUENCE 343 AA; 38949 MW; 894A870642B545DD CRC64;

Query Match 32.8%; Score 636; DB 1; Length 343;  
 Best Local Similarity 41.6%; Pred. No. 2,9e-35;  
 Matches 143; Conservative 58; Mismatches 115; Indels 28; Gaps 6;

QY 19 KEQFESYFGKRIAVDASMSIYQFLVVGRTGNETLNEGEVTSLLQGNFRIILLSA 78  
 D5 12 KEFLHNLGYCKRIALDNLAIYQFLSTIRORDTPLMDSGRITSLSLGFTYRLMLDA 71  
 QY 79 GIKPVYFEDGKPPDMKKOELAKRYSKRDATKDLTEAVENGDATREKISKRIYKTRGH 138  
 D5 72 GIKPVYFEDGKPPAFKKLEKREAREBEETRWKALAKGLEEKRAKQNRATKYNEK 131  
 QY 139 NEDCKRLRLIMGVYVAPSEAPAEACALCINDKRVAVASSEDMSLTFGAPRLRIIMP 198  
 D5 132 IDPAKKILLOMGPIYCAPSEGEQAAYVAGKGDVVASASQDVLISLFTPRJVRLLIT 191  
 QY 199 SSKKIPYMEVDV-----AVLEELITMQFTDLCGCCY-CISIKIGGOTL 248  
 D5 192 GKRRKPKGKDIYVIEKPELILEVELKELTKTRKILFIALVQTDYNGQIKGIGPKKAI 251  
 QY 249 KLIRHQSIESLENNKQ---RYQIPEDMPYQEARLKEPMTVDIFELIKATPADERG 305  
 D5 252 EIVK-----YSKPLAFQROSDVLYAIKEFFLNPTTDY-SLKRRPDEER 299  
 QY 306 LISHLYKNGFNEDRYVKAIEKIKSAKNNSSQGRLESF---KP 346  
 D5 300 IIRFLCEHDFSEPRVNGLEFLKAIKAGKQSTLESWFKKKP 343

RESULT 2  
 FEI\_PYRO STANDARD: PPT; 343 AA.  
 AC 050123;  
 DT 26-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Flap structure-specific endonuclease (EC 3.1.1.1).  
 GN Flap OR P1415.  
 OS Pyrococcus horikoshii.  
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
 OC Pyrococcus.  
 OX NCBI\_TaxID:53953;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-OT3;  
 RX MEDLINE:98344137; PubMed:9679194;  
 RA Kawarabayashi Y., Sawada M., Horikawa H., Halkawa Y., Hino Y.,

RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyara A., Nagai Y.,  
 RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,  
 RA Furunashi T., Tanaka T., Kudo Y., Yamazaki J., Katsuda N., Gyuichi A.,  
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb P.T., Horikoshi K.,  
 RA Masuchi Y., Shizuya H., Kikuchi H.;  
 RA \*Complete sequence and gene organization of the genome of a hyper-  
 RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";  
 RL DNA Res. 5:55-76(1998).

CC FUNCTION: Endonuclease that cleave the 5' overhanging flap  
 CC structure that is generated by displacement synthesis when DNA  
 CC polymerase encounters the 5' end of a downstream Okazaki fragment.  
 CC Has 5' endo-/exonuclease and 5' pseudo-Y-endonuclease activities.  
 CC Cleaves the junction between single and double-stranded regions of  
 CC flap DNA (By similarity).

CC COFACTOR: Binds 2 magnesium ions per subunit (By similarity).  
 CC - SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. FPK1  
 CC SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaborat  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstati  
 CC the European Bioinformatics Institute. There are no restrictions on  
 CC use by non-profit institutions as long as its content is in no  
 CC modified and this statement is not removed. Usage by and for comm  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/annou>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AP000006; HAA30521.1; ..  
 DR PIR: A71015; A71015.  
 DR PDB: 1MC8; 16-OCT-02.  
 DR HAMAP: MF\_00614; .. 1.  
 DR InterPro: IPR000513; Exo\_N.1.  
 DR InterPro: IPR031584; HH2\_2.  
 DR InterPro: IPR006086; XPG\_1.  
 DR InterPro: IPR006085; XPG\_N.  
 DR InterPro: IPR006084; XPG\_Rad.  
 DR Pfam: PF00867; XPG\_1; 1.  
 DR Pfam: PF00752; XPG\_N; 1.  
 DR PRINTS: PR00853; XPGRADSUPER.  
 DR SMART: SM00279; HH2; 1.  
 DR SMART: SM00484; XPG1; 1.  
 DR SMART: SM00485; XPGN; 1.  
 DR PROSITE: PS00841; XPG\_1; FALSE NEG.  
 KM Hydrolyase; Nuclease; Endonuclease; Magnesium; Metal-binding;  
 KW Complete proteome; 3D-structure.  
 FI METAL 154  
 SI SEQUENCE 343 AA; 38947 MW; 8BFD025F732C1118 CRC64;

Query Match 32.8%; Score 642.5; DB 1; Length 343;  
 Best Local Similarity 39.2%; Pred. No. 5e-35;  
 Matches 142; Conservative 66; Mismatches 119; Indels 35; Gaps

QY 1 MCQIKLITLADNAKAKKQKQFESYFGKRIAVDASMSIYQFLVVGRTGNETLNEGE 6  
 D5 1 MGVY ---TGDIVR--KEIDLENLGYCKRIALDNLAIYQFLSTIRORDTPLMDSGR 5  
 QY 61 VTSILQSEFNKTRILFAGIKPVYFEDGKPPDMKKOELAKRYSKRDATKDLTEAVENGD 1  
 D5 54 IISHSLSEYFTINLHFAQIKFAVYFEDGKPPDMKKOELAKRYSKRDATKDLTEAVENGD 1  
 QY 121 KDAIEKLSKPIYKVRQINEDCKRLRLIMGVYVAPSEAPAEACALCINDKRVAVASSED 1  
 D5 114 LEAKRYVQRATKYNEKIDPAKKILLOMGPIYCAPSEGEQAAYVAGKGDVVASASQDVL 1  
 QY 191 MCSIIFDAPRFLIMDSSEKRIYMEVDV-----AVLEELITMQFTDLCGCCY-CISIKIG 2  
 D5 234 YDSILTFGAPRLRIIMPIDPAKKILLOMGPIYCAPSEGEQAAYVAGKGDVVASASQDVL 2  
 QY 232 GDSY-CISIKIGGGOTATKLIRHQSIESLENNKQ---RYQIPEDMPYQEARLKEPMTVD 2  
 D5 234 GTDYNPQGVKCIQPKALEIVR-----YSRPLAFQROSDVLYAIKEFFLNPTTDY-SLK 2  
 QY 288 NVILDIPELKATPADERGISFLVYKNGFNEDRYVKAIEKIKSAKNNSSQGRLESF--- 3  
 D5 300 IIRFLCEHDFSEPRVNGLEFLKAIKAGKQSTLESWFKKKP 343







```

D1 28-FEB-2003 (rel. 41, last annotation update)
DE Flap structure-specific endonuclease (EC 3.-.-.-).
CN ENR OR S10210.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=111955.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RX MEDLINE=21456156; PubMed=11572475;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Aukai A., Kosugi R., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takayama M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudo Y., Yamazaki J., Kushiya N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yamagi M., Nishimura M., Yamagishi A.,
RA Ohama T., Kikuchi H.
RT Complete genome sequence of an aerobic thermophilic
RI Crenarchaeon, Sulfolobus tokodaii strain7.
RL DNA Res. 8:123-140(2001).
CC -1- FUNCTION: Endonuclease that cleave the 5'overhanging flap
CC structure that is generated by displacement synthesis when DNA
CC polymerase encounters the 5' end of a downstream Okazaki fragment.
CC Has 5' endo-/exonuclease and 5' pseudo-Y-endonuclease activities.
CC Cleaves the junction between single and double-stranded regions of
CC flap DNA (by similarity).
CC -1- COFACTOR: Binds 2 magnesium ions per subunit (by similarity).
CC -1- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. FEN1
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collabora-
CC tion between the Swiss Institute of Bioinformatics and the EMBL outstat-
CC ion at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AP000981; BAB5171.1; ALT_INIT.
DR HAMAP: MF_00614; 1.
DR Interpro: IPR000513; EXO_N_1.
DR Interpro: IPR003584; HNH_2.
DR Interpro: IPR006085; XPG_1.
DR Interpro: IPR006085; XPG_N.
DR Interpro: IPR006084; XPGC_Rad.
DR Pfam: PF00867; XPG_1; 1.
DR Pfam: PF00752; XPG_N; 1.
DR PRINTS: PR00853; XPGRADSUPER.
DR SMART: SM00273; HNH2; 1.
DR SMART: SM00484; XPG1; 1.
DR PROSITE: PS00841; XPG_1; FALSE_NEG.
KW Hydrolyase; Nuclease; Endonuclease; Magnesium; Metal-binding;
KW Complete proteome.
FT METAL 156 156 MAGNESIUM 1 (BY SIMILARITY).
FT SEQUENCE 351 AA; 39656 MW; DC48398E1CA06C4 CRC64.

```

```

Query Match 30.0%; Score 582; DB 1; Length 351;
Best Local Similarity 37.9%; Pred. No. 1; Le-21;
Matches 135; Conservative 73; Mismatches 120; Indels 28; Gaps 7;

```

```

QY 10 LADNAPKAMEQKQFSTFGKRIAYDASMSYQFLIYVGRGMEFLINAGEVISHQCF 65
DB 3 LAELVEIEIKKELSPAEIKGRKISIDVYNAFYQFLAIRTQDGTPLMDSQGRVISHLNGE 62
QY 70 NRTILLEGZIKPVYFDGKPPDKKQFLAKRYSKRDCAKDTLQTEAVEVD--KDAIEK 127
DB 11:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 122
QY 63 YRTISILEEGLIPIYVDGKPPQKQKQELERKKRVVEEAKKRIEOKATREKISTSELKRY 122
DB 128 SKRCVYTRQHNEDCKRLRLMGVPVVEAESEAEACALCINDKFAVASELMSSTETG 187
DB 123 AQMSIRLTINEMARESKELLKAMGIPVQCAESEAEAMVYINTGLSMATASQYDSELDF 182
QY 188 APFRLHIMPSSKKT-----VAFPIVAVYENLILTMQFIDLCILGQNY- 235

```

```

DB 193 AKR::RNLTSGKRKLPKQDYVEIKRPLDLDL--LTKRLGLTRQGLDGLIIVGIDYN
QY 236 CDSIKIGSGQTALKLIRQSGSTLENLNDKRYQIE--DWPQZARRLFKSPNVLTD
DB 241 PDGKIGYGVKTAIRIKKYSLEKALE-----KQETPRKYNFVWFRLSLFKPOVEP
QY 293 LPELKNTPDDEFGILSPLYKNGKNEEDVTAIKKISAKK-----SSQGLRSLSF 344
DB 296 KENIEDVCDNSKRLDILVKTQHPDNEERKNGIKERLEKAKKAEAGASQTLQDQWF 351

RESULT 12
FEN1_SITSO
ID FEN1_SITSO STANDARD: PKT: 349 AA.
AC Q98003.
DI 28-FEB-2003 (rel. 41, last sequence update)
DI 28-FEB-2003 (rel. 41, last annotation update)
DE Flap structure-specific endonuclease (EC 3.-.-.-).
CN ENR OR S502179.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=22871.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35062 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA Ste O., Singh R.K., Contatoulet F., Zivanovic Y., Allard G.,
RA Awaer M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Erasmo G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doctolite W.F., Eugenet M., Gasteirland T.,
RA Garrett R.A., Ragan M.A., Sengen C.W., Van der coast J.J.
RI "The complete genome of the crenarchaeon Sulfolobus solfataricus P2."
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
CC -1- FUNCTION: Endonuclease that cleave the 5'overhanging flap
CC structure that is generated by displacement synthesis when DNA
CC polymerase encounters the 5' end of a downstream Okazaki fragment.
CC Has 5' endo-/exonuclease and 5' pseudo-Y-endonuclease activities.
CC Cleaves the junction between single and double-stranded regions o
CC flap DNA (by similarity).
CC -1- COFACTOR: Binds 2 magnesium ions per subunit (by similarity).
CC -1- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. FEN1
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collabora-
CC tion between the Swiss Institute of Bioinformatics and the EMBL outstat-
CC ion at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commo
CC entities requires a license agreement (see http://www.isb-sib.ch/anno
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE006556; AAK40525.1; ALT_INIT.
DR HAMAP: MF_00614; 1.
DR Interpro: IPR000513; EXO_N_1.
DR Interpro: IPR003584; HNH_2.
DR Interpro: IPR006085; XPG_1.
DR Interpro: IPR006085; XPG_N.
DR Interpro: IPR006084; XPGC_Rad.
DR Pfam: PF00867; XPG_1; 1.
DR Pfam: PF00752; XPG_N; 1.
DR PRINTS: PR00853; XPGRADSUPER.
DR SMART: SM00273; HNH2; 1.
DR SMART: SM00484; XPG1; 1.
DR PROSITE: PS00841; XPG_1; FALSE_NEG.
KW Hydrolyase; Nuclease; Endonuclease; Magnesium; Metal-binding;
KW Complete proteome.
FT METAL 154 154 MAGNESIUM 1 (BY SIMILARITY).
FT SEQUENCE 349 AA; 39497 MW; 721C3455D749C946 CRC64.

```

Query Match 29.0% Score 562: DB 1: Length 349;  
Best Local Similarity 36.1%: Pred. No. 2,4e-30;  
Matches 130: Conservative 69: Mismatches 123: Indels 38: Gaps 6;  
QY 10 LADNPKRAMEQKESTFGKRIADVANSITVPLIVGRIGMETLTNEAGFVSHIQGMF 69  
DB 3 LADLVKDKVRELSESELKRSVIDGYNALYQFZAAIFQPDGTPDKMSQGVSHLSGLF 62  
QY 70 NRTTLLFAGIKIPVYFPGKPPDKKQFIARYSKRDATDOLTENVGVCKKAFIKISK 129  
DB 63 YRTINILEESVPIVPEKPPKSELEERRRRAKEAEKRLRAVSEGIIEELIKYSQ 122  
QY 130 RTVYTRQHNEDCKRLKMGVPEVAPSEAEACALCIDKVFAYASECMDSLTGAP 189  
DB 123 ALIRISNIMVESKLLRANCIPIVQANSEGEAEAAVINKIGLSMAASQVDAILLCAK 182  
QY 190 RFLRLMDPSSKRP-----VMEFDVAVLEELSLTMDQFIDLCIGCCGV-CD 237  
DB 183 RLVRNLITTRGRKRLPNKDYVEIKPELIEIEI-LLNKLGITRQDLIDIGLTIDVNP 240  
QY 238 SIKGIGOTAKLRLRQHSISILE-----NLNDRFQIPEDMPEYCARLFFKPN 288  
DB 241 GIRGGERALKTIKKYKIKEMEYGISKKDINFND-----EIRGLFLNFO 289  
QY 289 VTLDIPELKWTPDEGLISFLVADNGFNEDRVTKAIEK---IKSAKNSQGRLESFF 344  
DB 290 VVKEPEALDLPNEQEDILNILLVEHNSEBRVXNGIHLTKAIKAKGASRQGLDRMF 349  
RESULT 13  
FEN\_METJA  
ID FEN\_METJA STANDARD: PRT: 326 AA.  
AC Q58839;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Flap structure-specific endonuclease (EC 3.1.1.1).  
GN FEN CR MJ1444.  
OS Methanococcus jannaschii.  
OC Archaea: Furoyarchaeota: Methanococci: Methanococcales;  
OC Methanocaldococcales; Methanocaldococcus.  
OX NCBI\_TaxID=2193;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;  
RX MEDLINE=96337999; PubMed=8688087;  
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,  
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Belch C.L.,  
RA Overbeek R., Kirkness E.F., Weinstock K.C., Merrick J.M., Gocdek E.,  
RA Scott J.L., Geoghegan N.S.M., Welman J.P., Furrman J.M., Nguyen F.,  
RA Usterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
RA \*Complete genome sequence of the methanogenic archaeon, Methanococcus  
RA jannaschii";  
RA Science 273:1059-1073(1996).  
RN [2]  
RP CHARACTERIZATION.  
RX MEDLINE=96440432; PubMed=975572;  
RA Rao H.G., Rosenfeld A., Wetmur J.C.;  
RA \*Methanococcus jannaschii flap endonuclease: expression, purification,  
RA and substrate requirements";  
RA J. Bacteriol. 180:5406-5412(1998).  
RN [3]  
RP CHARACTERIZATION AND X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
RX MEDLINE=99200693; PubMed=10102570;  
RA Bae K.W., Bae K.W., Cho C.S., Hwang K.Y., Kim H.-K., Sung H.-C.,  
RA Cho Y.;  
RA "Expression, purification, characterization and crystallization of  
RA flap endonuclease-1 from Methanococcus jannaschii.";  
RA Mol. Cells 9:45-48(1999).

EN 141  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;  
RX MEDLINE=98363215; PubMed=969635;  
RA Huang K.Y., Bae K., Kim H.-Y., Cho Y.;  
RA \*The crystal structure of flap endonuclease-1 from Methanococcus  
RA jannaschii";  
RA Nat. Struct. Biol. 5:707-713(1998).  
CC -1- FUNCTION: Endonuclease that cleaves the 5' overhanging flap  
CC structure that is generated by displacement synthesis when DNA  
CC polymerase encounters the 5' end of a downstream Okazaki fragment.  
CC has 5' endo-/exonuclease and 5' pseudo-Y-endonuclease activities.  
CC cleaves the junction between single and double-stranded regions of  
CC flap DNA.  
CC -1- COFACTOR: Binds 2 magnesium ions per subunit.  
CC -1- MISCELLANEOUS: Its optimum pH is 6.7.  
CC -1- SIMILARITY: BESONOS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. FEN1  
CC SUBFAMILY.  
CC  
CC This SWISS-PROT entry is copyrighted. It is produced through a collabora-  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on  
CC use by non-profit institutions as long as its content is in no  
CC modified and this statement is not removed. Usage by and for commer-  
CC entities requires a license agreement (See <http://www.isb-sib.ch/annou>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL: U67585; AB999454.1;  
DR PIR: G64480; G64480.  
DR PDB: 1A76; 03-AUG-99.  
DR PDB: 1A77; 03-AUG-99.  
DR TIGR: MJ1444;  
DR HMAP: ME\_00614; 1.  
DR InterPro: IPR00513; EXO\_N1.  
DR InterPro: IPR003584; HHH\_2.  
DR InterPro: IPR006086; XPG\_1.  
DR InterPro: IPR006085; XPG\_N.  
DR InterPro: IPR005084; XPGC\_Rad.  
DR Pfam: PF00867; XPG\_1; 1.  
DR Pfam: PF00752; XPG\_N; 1.  
DR PRINTS: PR00853; XPGRADSUPER.  
DR SMART: SM00279; HNH2; 1.  
DR SMART: SM00484; XPG1; 1.  
DR SMART: SM03485; XPGN; 1.  
DR Hydrolase; Nuclease; Endonuclease; Magnesium; Metal-binding;  
KW 3D-structure; Complete proteome.  
FT METAL 154  
FT SEQUENCE 326 AA; 37527 MW; 311427FBA67580 CRC64;  
Query Match 28.9% Score 560.5: DB 1: Length 326;  
Best Local Similarity 36.0%: Pred. No. 2,4e-30;  
Matches 132: Conservative 71: Mismatches 121: Indels 23: Gaps  
QY 1 MGIGIKRLKADNPKRAMEQKESTFGKRIADVANSITVPLIVGRIGMETLTNEAGFVSHIQGMF 69  
DB 2 MGAV-----EGDIPKRLI--ISEELKGRKVALDGMKNAITVPLIKGSPLRNKGK 62  
QY 62 VTSHEGEMNPTIRLIFAGIKIPVYFPGKPPDKKQFIARYSKRDATDOLTENVGVCKKAFIKISK 129  
DB 64 ITSAYNVVFKTILHLKNDITPEVDFGPPKIKKK-RKYRREKAEAEKRLRAVSEGIIEELIKYSQ 122  
QY 122 KDALEKSKRTVYTRQHNEDCKRLKMGVPEVAPSEAEACALCIDKVFAYASECMDSLTGAP 189  
DB 114 PELAKTKAKSVITTRKRVNCKTILSIKQIPVFAVSEGEACASVYAKKGDVAVVSSD 182  
QY 181 MDSTIFAPRFLRLMDPSSKRPVMEFDVAVLEELSLTMDQFIDLCIGCCGV-CDST 237  
DB 174 YDALVYAPVVRHL-ETTKRPHL-IEHNPIVLEURLSDLDIDVAFMGDTYNGGV 240  
QY 240 KGIGIGOTAKLRLRQHSISILE-----NLNDRFQIPEDMPEYCARLFFKPN 288  
DB 232 KGIGIKRVAEYVRS-GVAKVLRK-----FVEYDEIKRIFKSPKV-DMYSLSLK 289

```

QY 300 APDEGLISFLKNGFNEDRYTAKLEIKS-AKKSSQGLSEFFK 345
Db 280 LPDEKSTIKELVDNFENFNRVKKHYDKLYLNANKTKOKTIDAMK 326

RESULT 14
FEN_METMA STANDARD: PRT: 338 AA.
AC 08PYF6:
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Flap structure-specific endonuclease (EC 3.1.1.1).
GN FEN OR MM0906.
OS Methanoscarchina mazel (Methanoscarchina trisla).
OC Archaea; Euryarchaeota; Euryarchaeota orders Incertae sedis;
OC Methanoscarchinales; Methanoscarchinaceae; Methanoscarchina.
OX NCBI_TaxID=2207;
HN 11
FP SEQUENCE FROM N.A.
RC STRAIN=JG01 / G01 / ATCC BAA-159 / DSM 3647 / CCM 88;
RX MEDLINE=2210827; PubMed 12125824;
RA Deppe-Mueller U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
RA Martinez-Arias R., Henne A., Wlezer A., Baerumt S., Jacob G.,
RA Bueggemann H., Lienard T., Christmann A., Boemcke M., Stecke R.,
RA Bhattacharya A., Lykilds A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA Fitts H.-J., Gottschalk G.;
RA "The genome of Methanoscarchina mazel: evidence for lateral gene
RA transfer between Bacteria and Archaea";
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
CC 1- FUNCTION: Endonuclease that cleave the 5'overhanging flap
CC structure that is generated by displacement synthesis when DNA
CC polymerase encounters the 5'end of a downstream Okazaki fragment.
CC Has 5'endo-/exonuclease and 5'pseudo-Y-endonuclease activities.
CC Cleaves the junction between single and double-stranded regions of
CC flap DNA (By similarity).
CC 1- COFACTOR: Binds 2 magnesium ions per subunit (By similarity).
CC 1- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. FEW1
CC SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE013116; AM30602.1;
CC HAMAP: MF_00614;
CC DR InterPro: IPR001986; EPSP_synthase.
CC DR InterPro: IPR000513; EXO_N1.
CC DR InterPro: IPR006086; XPG_N.
CC DR InterPro: IPR006085; XPG_N.
CC DR InterPro: IPR006084; XPG_N_Rad.
CC DR Pfam: PF00867; XPG_I_1.
CC DR Pfam: PF00752; XPG_N_1.
CC DR PRINTS: PR00834; XPGRADSUPER.
CC DR PROSITE: PS00841; XPG_1; FALSE_NEG.
CC KW Hydrolyase; Nuclease; Endonuclease; Magnesium; Metal-binding;
CC Complete proteome.
CC FT METAL 154 154 MAGNESIUM 1 (BY SIMILARITY).
CC SEQUENCE 338 AA; 37711 MW; 59069803B0F30E3 CRC64.

Query Match 26.9%; Score 500.5; DB 1; Length 338;
Best Local Similarity 37.3%; Pred. No. 2.9e-30;
Matches 126; Conservative 69; Mismatches 124; Indels 19; Gaps 5;

```

```

QY 16 MKEQKFE--SYFGKIAVDASMSITQPLIVORTGNETITNEAGVSHIQQNMRIRL 75
Db 9 LQKRIKELSDLSNRVAVADANTLHQLFSLTIQKRGSPVANSQCKVISHLSGLIYFASL 68
QY 76 LKNGIKPVIVFGDKPKDKKXDELANRYSKRQDAIKDLTAEVVGDKAIEKLSRYKYVT 135

RESULT 15
FEN_METMA STANDARD: PRT: 328 AA.
AC 027670:
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Flap structure-specific endonuclease (EC 3.1.1.1).
GN FEN OR MH1633.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteriales;
OC Methanobacteriaceae; Methanobacter.
OX NCBI_TaxID=187420;
HN 11
FP SEQUENCE FROM N.A.
RC STRAIN=Deila H;
RX MEDLINE=98037514; PubMed:9371463;
RA Smith D.R., Deusselle-Stamm L.A., Belonghery G., Lee H.-M., Dupois J.,
RA Adrege T., Hashizadeh R., Hlaekily D., Cook R., Gilbert K.,
RA Harrison D., Huang L., Kung'u P., Linn M., Rothner R., Gu D.,
RA Shadforth R., Vicario R., Wang Y., Wierzbowski J., Gibson R.,
RA Zeman N., Zarso A., Bush D., Sater N., Patwell L., Frubakar S.,
RA McDougall S., Shiner G., Goyal A., Petrovski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling C., Reeve J.N.;
RA "Complete genome sequence of Methanobacterium thermoautotrophicum
RA deltaH: functional analysis and comparative genomics";
RL J. Bacteriol. 179:7135-7153(1997).
CC 1- FUNCTION: Endonuclease that cleave the 5'overhanging flap
CC structure that is generated by displacement synthesis when DNA
CC polymerase encounters the 5'end of a downstream Okazaki fragment.
CC Has 5'endo-/exonuclease and 5'pseudo-Y-endonuclease activities.
CC Cleaves the junction between single and double-stranded regions of
CC flap DNA (By similarity).
CC 1- COFACTOR: Binds 2 magnesium ions per subunit (By similarity).
CC 1- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. FEW1
CC SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collabora
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commerc
CC entities requires a license agreement (See http://www.isb-sib.ch/annou
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE009222; AAB86106.1;
CC FTR: C69085; C69085.
CC DR HAMAP: MF_00614;
CC DR InterPro: IPR000513; EXO_N1.
CC DR InterPro: IPR003584; XPG_N2.
CC DR InterPro: IPR006086; XPG_I_1.
CC DR InterPro: IPR006085; XPG_N.

```

DR InterPro: IPRO06054; XPG\_Rad.  
DR Pfam: PF00867; XPG\_L1.  
DR Pfam: PF00752; XPG\_N1.  
DR Pfam: PF00752; XPG\_N2.  
DR Pfam: PF00853; XPRADSCP.  
DR SMART: SM00279; Hh2; 1.  
DR SMART: SM00484; XPG1; 1.  
DR SMART: SM00485; XPGN; 1.  
DR PROSITE: PS00841; XGC; 1; FALSE\_NEG.  
DR Hydroxylase, Nuclease, Endonuclease; Metal-Binding; Complete proteome.  
FT METAL 154 MAGNESIUM 1 (ZY SIMILARITY).  
SQ SEQUENCE 328 AA: 37170 MW: CACICG04551FDEA CFC64;



DR PROSITE: PS00841: XPG\_1: 1  
 SU SEQUENCE 368 AA: 42792 MW: F0146AFA55A7364 CRC64:  
 Query Match 93.2%; Score 1807.5; DB 10; Length 360;  
 Best Local Similarity 92.1%; Pred. No. 3,2e-124;  
 Matches 350; Conservative 16; Mismatches 13; Indels 1; Gaps 1;

QY 1 MSIKGLTKLADNMPKAKKEGFESYFGKRIAVASASTYQFLIVSGRTGMEILLTNEAGE 60  
 |||||||  
 DB 1 MGKGLTKLADNMPKAKKEGFESYFGKRIAVASASTYQFLIVSGRTGMEILLTNEAGE 60  
 QY 61 VTSHLQGMENRTIRLLENGIKPVYVFDGKPPDMKKOELAKRYSKRDATKQTEAEVGG 120  
 |||||||  
 DB 61 VTSHLQGMENRTIRLLENGIKPVYVFDGKPPDMKKOELAKRYSKRDATKQTEAEVGG 120  
 QY 121 KDAIEKSKRYKYVTRQNEQCKRLRLMGVPVVEADSEAEACALCIDKYEVAASED 180  
 |||||||  
 DB 121 KDAIEKSKRYKYVTRQNEQCKRLRLMGVPVVEADSEAEACALCIDKYEVAASED 180  
 QY 181 MDLTFGAPRFLRLMDPSSKKIPMEFDVAKVLELELTMDQFDIDCIGGCDYCDSTK 240  
 |||||||  
 DB 181 MDLTFGAPRFLRLMDPSSKKIPMEFDVAKVLELELTMDQFDIDCIGGCDYCDSTK 240  
 QY 241 GIGGOTALKLRHOSISILENLKORYQIPEDMPYQBARLRFKENVTLDIPELKRTA 300  
 |||||||  
 DB 241 GIGGOTALKLRHOSISILENLKORYQIPEDMPYQBARLRFKENVTLDIPELKRTA 300  
 QY 301 PDEGGLISFLYKDNQNFEDRYTKATEIKISAKNKSOGRCFSEFKPATISAPLKPKETS 360  
 |||||||  
 DB 301 PDEGGLISFLYKDNQNFEDRYTKATEIKISAKNKSOGRCFSEFKPATISAPLKPKETS 360  
 QY 361 DKTSKAANKKTK-AGSKK 379  
 |||||||  
 DB 361 DKTSKAANKKTK-AGSKK 380

RESULT 2  
 ID 065251 PRELIMINARY: PRT: 362 AA.  
 AC 065251;  
 DT 01-AUG-1998 (TRENHIREL: 07, Created)  
 DT 01-AUG-1998 (TRENHIREL: 07, Last sequence update)  
 DT 01-OCT-2002 (TRENHIREL: 22, Last annotation update)  
 DE F2IE0.3 protein.  
 GN F2IE0.3  
 OS Arabidopsis thaliana (Mouse-ear cress);  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Eudicotyledons; core eudicots; Fossidae;  
 OC Eurosidia II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 KA Davidson S., Kolling T., David M., O'Brien D.;  
 RT The sequence of A. thaliana F2IE0.3;  
 RL Submitted (Apr-1998) to the EMBL/GenBank/DBJ databases.  
 RN 12  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 KA Submitted (Apr-1998) to the EMBL/GenBank/DBJ databases.  
 RN 13  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 KA Waterston R.;  
 RL Submitted (Apr-1998) to the EMBL/GenBank/DBJ databases.  
 RN 14  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 KA Wilson R.;  
 RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF058914: AAC13596.1: -

DR HSRP: Q58839: 1A76.  
 DR InterPro: IPR02421: 5\_3\_exonuclease.  
 DR InterPro: IPR000513: Exo\_N1.  
 DR InterPro: IPR003584: HHR\_2.  
 DR InterPro: IPR006884: XPGC\_Pad.  
 DR InterPro: IPR006086: XPG\_I.  
 DR InterPro: IPR006093: XPG\_N.  
 DR Pfam: PF01957: 5\_3\_exonuclease\_1.  
 DR Pfam: PF00857: XPG\_1.  
 DR Pfam: PF00752: XPG\_N; 1.  
 DR PRINTS: PR00853: XPGRADSUPER.  
 DR SMART: SM00279: HHR\_2; 1.  
 DR SMART: SM00484: XPGC; 1.  
 DR SMART: SM00485: XPGN; 1.  
 DR PROSITE: PS00841: XPG\_1: 1.  
 SU SEQUENCE 362 AA: 41205 MW: A4DD1706C3AD250 CRC64:  
 Query Match 67.7%; Score 1313.5; DB 10; Length 362;  
 Best Local Similarity 72.1%; Pred. No. 4,2e-88;  
 Matches 259; Conservative 30; Mismatches 33; Indels 37; Gaps

QY 1 MGKGLTKLADNMPKAKKEGFESYFGKRIAVASASTYQFLIVSGRTGMEILLTNEAGE 60  
 |||||||  
 DB 1 MGKGLTKLADNMPKAKKEGFESYFGKRIAVASASTYQFLIVSGRTGMEILLTNEAGE 60  
 QY 61 VTS.....HLQGMENRTIRLLENGIKPVYVFDGKPPDMKKOELAKRYSKRDATKQTEAEVGG 120  
 |||||||  
 DB 61 VTS.....HLQGMENRTIRLLENGIKPVYVFDGKPPDMKKOELAKRYSKRDATKQTEAEVGG 120  
 QY 120 DKKPEELKROLARYSKRADATADLTGALVACNKEITFYKSKRYKVTRQNEQCKRLRL 180  
 |||||||  
 DB 120 DKKPEELKROLARYSKRADATADLTGALVACNKEITFYKSKRYKVTRQNEQCKRLRL 180  
 QY 180 RLMGVPVVEADSEAEACALCIDKYEVAASEDMDLTFGAPRFLRLMDPSSKKIPMEFDVAKVLELELTMDQFDIDCIGGCDYCDSTK 240  
 |||||||  
 DB 180 RLMGVPVVEADSEAEACALCIDKYEVAASEDMDLTFGAPRFLRLMDPSSKKIPMEFDVAKVLELELTMDQFDIDCIGGCDYCDSTK 240  
 QY 240 FFEVAKRIEFLQTLMDQFDIDCIGGCDYCDSTKIGGOTALKLRHOSISILENLK 300  
 |||||||  
 DB 240 FFEVAKRIEFLQTLMDQFDIDCIGGCDYCDSTKIGGOTALKLRHOSISILENLK 300  
 QY 300 EYQVTFEPMYQBARLRFKENVTLDIPHL-  
 |||||||  
 DB 300 EYQVTFEPMYQBARLRFKENVTLDIPHL-  
 |||||||

RESULT 3  
 ID 070043 PRELIMINARY: PRT: 362 AA.  
 AC 070043;  
 DT 01-FEB-1997 (TRENHIREL: 02, Created)  
 DT 01-FEB-1997 (TRENHIREL: 02, Last sequence update)  
 DT 01-OCT-2002 (TRENHIREL: 22, Last annotation update)  
 DE 5' nuclease XEN-A.  
 GN XEN-A.  
 OS Xenopus laevis (African clawed frog);  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8359;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RC BIBIKOVA M., CHI E., WU B., KIM K., GARROD J.;  
 KA Submitted (Jul-1996) to the EMBL/GenBank/DBJ databases.  
 RN 12  
 RP SEQUENCE FROM N.A.  
 RC LITTELL, COX L.S.;  
 KA Cloning and investigation of Xenopus laevis developmental expression  
 and function in DNA replication;  
 RT Submitted (May-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U64563: AA06676.1: -  
 DR EMBL: AF053397: AA02814.1: -  
 DR HSRP: Q58839: 1A76.

DR InterPro: IPR002421; 5\_3-exonuclease.  
 DR InterPro: IPR000513; Exo\_N.1.  
 DR InterPro: IPR003584; HNH\_2.  
 DR InterPro: IPR006084; XPG\_Rad.  
 DR InterPro: IPR006085; XPG\_I.  
 DR InterPro: IPR006086; XPG\_N.  
 DR Pfam: PF01367; 5\_3-exonuclease; 1.  
 DR Pfam: PF00867; XPG\_1; 1.  
 DR Pfam: PF00752; XPG\_N; 1.  
 DR PRINTS: PR00853; XPGRADSUPER.  
 DR SMART: SMO0279; HNH2; 1.  
 DR SMART: SMO0484; XPG1; 1.  
 DR SMART: SMO0485; XPGN; 1.  
 DR Endonuclease.  
 DR SEQUENCE 382 AA: 42668 MW: 98180EDAD158D57 CRC64:

Query Match 55.1%; Score 1069; DR 13; Length 382;  
 Best Local Similarity 55.7%; Pred. No. 3,6e-70;  
 Matches 210; Conservative 63; Mismatches 98; Indels 6; Gaps 4;

QY 1 MGKIGTLKLLADNAPKAKKEQKFESEYFGKILAVDASMSIYQFLIVRTGMLTNEAGE 60  
 D 1 MGIHGLAKLADVAPALKEHDKSYFGKAVVADSMCTYQFLAVRQDQ-NITQNEAGE 59  
 QY 61 VTSILQGFENRTIRLEAGIKPVYVFDCKPPPMKQGLAKRYSKDGLTDLTEAVEGQ 120  
 D 60 TSHLMGMFYRTIRBLEHGIRKPVYVFDCKPPPMKSGELAKSRRAAEAKLLEAAEEAGE 119  
 QY 121 KDAIEKLSKRYKVTQHNEDCKRLRLMGVNVVAPSEAECAALCINDKRVAVASED 180  
 D 120 VENIEKTRKLVKVKYKQHNEDCKRLTLMGITVYVADPEAEKTCALVAKKVVAAAEED 179  
 QY 181 MSLTFGARPLRLHLMDPSKKCPVMEFDVAKVLELELTMDQFLDGLTCCGYGCEIK 240  
 D 180 MDALTFGTIVLLRLHLTASAKKLPQIFHLNIVDIDGITHQFVLDLILSSYGCETIR 239  
 QY 241 GIGGQIALKLIROHGSIESLENLKKDRYQIPEDMPYQEARLKEPNVT-LDIPPELKW 299  
 D 240 GIGPKRAIDLIRQHKIDIELINIDIKKYPFNNLHKEAKLLEPEVVTDLFLKWI 298  
 QY 300 APDEEGISFLVYKDNFNEEDRYTKATEIKTSKNNSSQCKLESFEPKATATSAF-KRKET 359  
 D 300 EPDEEGIVAFMCGEKQFSEDRIRNKAKKLAKKNGSTQGRDIDFVKVGSVSS-TKKREA 358  
 QY 360 SEKTSKAANKKTKAGG 376  
 D 359 E---SKGSAKKRAKTGG 372

## RESULT 4

ID P70054 PRELIMINARY; PRI: 382 AA.  
 AC P70054;  
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE XENB.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 CC NCBI\_TaxID=8355;  
 CX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Head;  
 RA MEDLIN:99059415; PubMed=9852084;  
 RA Bidikova M., Wu B., Chl E., Kim K.H., Trautman J.K., Carroll D.;  
 RT "Characterization of FEN-1 from Xenopus laevis. cDNA cloning and role  
 in DNA metabolism.";  
 RL J. Biol. Chem. 273:34222-34229(1998).  
 DR EMBL: U63141; AAB08478.1;  
 DR HSP: Q58833; I476.  
 DR InterPro: IPR002421; 5\_3-exonuclease.  
 DR InterPro: IPR002421; 5\_3-exonuclease.

DR InterPro: IPR000513; Exo\_N.1.  
 DR InterPro: IPR003584; HNH\_2.  
 DR InterPro: IPR006084; XPG\_Rad.  
 DR InterPro: IPR006085; XPG\_I.  
 DR InterPro: IPR006086; XPG\_N.  
 DR Pfam: PF01367; 5\_3-exonuclease; 1.  
 DR Pfam: PF00867; XPG\_1; 1.  
 DR Pfam: PF00752; XPG\_N; 1.  
 DR PRINTS: PR00853; XPGRADSUPER.  
 DR SMART: SMO0279; HNH2; 1.  
 DR SMART: SMO0484; XPG1; 1.  
 DR SMART: SMO0485; XPGN; 1.  
 DR SEQUENCE 382 AA: 42865 MW: 10548916A2324460 CRC64:

Query Match 55.1%; Score 1068; DR 13; Length 382;  
 Best Local Similarity 56.2%; Pred. No. 4.2e-70;  
 Matches 222; Conservative 60; Mismatches 99; Indels 6; Gaps 4;

QY 1 MGKIGTLKLLADNAPKAKKEQKFESEYFGKILAVDASMSIYQFLIVRTGMLTNEAGE 60  
 D 1 MGIHGLAKLADVAPALKEHDKSYFGKAVVADSMCTYQFLAVRQDQ-NITQNEAGE 59  
 QY 61 VTSILQGFENRTIRLEAGIKPVYVFDCKPPPMKQGLAKRYSKDGLTDLTEAVEGQ 120  
 D 60 TSHLMGMFYRTIRBLEHGIRKPVYVFDCKPPPMKSGELAKSRRAAEAKLLEAAEEAGE 119  
 QY 121 KDAIEKLSKRYKVTQHNEDCKRLRLMGVNVVAPSEAECAALCINDKRVAVASED 180  
 D 120 VENIEKTRKLVKVKYKQHNEDCKRLTLMGITVYVADPEAEKTCALVAKKVVAAAEED 179  
 QY 181 MSLTFGARPLRLHLMDPSKKCPVMEFDVAKVLELELTMDQFLDGLTCCGYGCEIK 240  
 D 180 MDALTFGTIVLLRLHLTASAKKLPQIFHLNIVDIDGITHQFVLDLILSSYGCETIR 239  
 QY 241 GIGGQIALKLIROHGSIESLENLKKDRYQIPEDMPYQEARLKEPNVT-LDIPPELKW 299  
 D 240 GIGPKRAIDLIRQHKIDIELINIDIKKYPFNNLHKEAKLLEPEVVTDLFLKWI 298  
 QY 300 APDEEGISFLVYKDNFNEEDRYTKATEIKTSKNNSSQCKLESFEPKATATSAF-KRKET 359  
 D 300 EPDEEGIVAFMCGEKQFSEDRIRNKAKKLAKKNGSTQGRDIDFVKVGSVSS-TKKREA 358  
 QY 360 SEKTSKAANKKTKAGG 376  
 D 359 E---SKGSKRSKTSKG 372

## RESULT 5

ID O57351 PRELIMINARY; PRI: 382 AA.  
 AC O57351;  
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Flap endonuclease 1.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 CC NCBI\_TaxID=8355;  
 CX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE:98204872; PubMed=9855864;  
 RA Kim K., Biade S., Matsumoto Y.;  
 RT "Involvement of flap endonuclease 1 in base excision DNA repair.";  
 RL J. Biol. Chem. 273:8542-8548(1998).  
 DR EMBL: AF036327; AAB88707.1;  
 DR HSP: Q58833; I476.  
 DR InterPro: IPR002421; 5\_3-exonuclease.  
 DR InterPro: IPR000513; Exo\_N.1.  
 DR InterPro: IPR003584; HNH\_2.  
 DR InterPro: IPR006084; XPG\_Rad.





DR EMBL: AK078778; BAC37390.1; -.  
DR MGD: MG1:102779; Fent1.  
DR InterPro: IPR002421; 5\_3-exonuclease.  
DR InterPro: IPR000513; Exo-N\_I.  
DR InterPro: IPR003584; HNH\_2.  
DR InterPro: IPR006084; XPGC\_Rad.  
DR InterPro: IPR006086; XPG-I.  
DR InterPro: IPR006085; XPG-N.  
DR Pfam: PF01367; 5\_3-exonuclease; 1.  
DR Pfam: PF00867; XPG-I; 1.  
DR Pfam: PF00752; XPG-N; 1.  
DR PRINTS: PRC0853; XPGRADSUPER.  
DR SMART: SM00279; Hnh2; 1.  
DR SMART: SM00484; XPG1; 1.  
DR SMART: SM00485; XPGN; 1.  
DR PROSITE: PS00841; XPG-1; 1.  
DR PROSITE: PS00842; XPG-2; 1.  
DR Pfam: Pfam00000; Exonuclease.  
DR SOURCE: 350 AA; 42623 MW; 1BF9031288BA6520D CRC64;  
Q

Query Match 53.8%; Score 1043; DB 11; Length 380;  
 Best Local Similarity 53.8%; Pred. No. 2.8e-68;  
 Matches 205; Conservative 69; Mismatches 99; Indels 8; Gaps 5

QY	1	MGIKGLTKLLADNAPRAKMEQKFESEYFQKRLAIVDASMSITQFLIVGRTQMSTLTHACE	60
Dh	2	MGTHIGLAK--LAAVAAASAIARENDIKSYFGRKVAIDASMSICYFLAAV--RQGGDYLVQWENGE	59
QY	61	VTSHLQGMENPTIRLLEAGIKPVVYVEQCKPDMQKQELARYSKRDQATDULLEAVENQD	120
Dh	60	TIHLQMGKRYRIRIRRMENQ--KRPVYVLCCKPQULQSGT--LARKSERRAAEKQIQAQVAFQM	119
QY	121	KDAEILSKRIYAVTRQHNFEQCK--LRLMGVPVVEAPEASEAFCAIQLNDKPYAVASD	160
Dh	120	EEVEKFTRLRYKAVYKQHNDECKHLLSLMGIPYILDAPASEAASQALAKKQKYYAANTD	159
QY	181	MDSLFPGADPRLRLHNDPSSKTIYMEEDVAKLEELTLTMDCTFDIHLCTG--QVQDQSK	240
Dh	180	MDCLTFGSEVLRLHNLASFAKRLPIQEPHLRLSRVLQETGLNQEOPVD--CILGSDQESFR	239
QY	241	GIGCGFALMLIHQHSIELEINLKKQRYQIRPDMQPOEARLTKENYTLD--ITELKM	298
Dh	240	GIGPKRAVLDLQKHSSIEIYLRDLDPSSYPAPEWML--KEAQOLFLEFV--LQDESEVLCW	298
QY	299	TAPDEGLISFLVQKDNQGFEDRVTKALEIKISAKNKSQGLSEFSPKATITAP--LRKE	358
Dh	299	SEPMNEELVFMGEGEOPSEERIRSGVAKRLSKNSQGSTQGRLDQDFEIVTISLSS--AKKH	357
QY	359	TSQKTSKAANKKTKYAGGKK	379
Dh	358	PE--PRGPAAKKAKATGAGK	377

```

RESULT: 8
08C5X6
ID 08C5X6          PREDIMINARY:      PRT:    411 AA.
AC 08C5X6;
DT 01-MAR-2003 (iEMBLrel. 23, Created)
DT 01-MAR-2003 (iEMBLrel. 23, last sequence update)
DT 01-MAR-2003 (iEMBLrel. 23, last annotation update)
DF Flap structure specific endonuclease 1.
CS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI-TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466551;
RA The FAWN Consortium.
RA The FAWN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
  62,770 full-length cDNAs."

```

KL. Nature 420:563-573(2002).  
 DK EMHL AK076977; BAC35544.1. -  
 SQ SEQUENCE 411 AA: 45619 MW: 007B331220454032D CRC64:

Query Match	53.6%	Score 1043	DB 11	Length 411
Best Local Similarity	53.8%	Pred. No. 3.2e-68		
Matches 205	Conservative 69	Mismatches 99	Indels 6	

[illegible]

```

RESULT 9
OR0069
ID OR0069 PRELIMINARY: PRT: 340 AA.
AC OR0069:
DT 01-JUN-2002 (TREMBL: 21, Created)
DT 01-JUN-2002 (TREMBL: 21, Last sequence update)
DT 01-MAR-2003 (TREMBL: 23, Last annotation update)
DE Hypothetical 42.6 kDa protein.
GS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CN NCBI_taxid:10090.
CX 11)
RN PP SEQUENCE FROM N.A.
RA Strausberg R.;
KI Submitted (APR-2002) to the EMBL/GenBank/DDRJ databases.
DR EMBL: BC027295; AAM27295.1; -.
DR InterPro: IPR000513; EXO_N.1.
DR InterPro: IPR005584; HEH_2.
DR InterPro: IPR006286; XPG_1.
DR InterPro: IPR006085; XPG_N.
DR Pfam: PF00867; XPG_L1; 1.
DR Pfam: PF00752; XPG_N; 1.
DR SMART: SMC0279; HHR2; 1.
DR SMART: SMC0484; XPG1; 1.
DR SMART: SMC0485; XPGN; 1.
DR PROSITE: PS00841; XPG_1; 1.
DR PROSITE: PS00842; XPG_2; 1.
KW Hypothetical protein.
SD SEQUENCE 340 AA: CRR35745358P55A7.CRC34:

```

Query Match	51.7%	Score	10427	DB 11	Length	3603			
Best Local Similarity	53.5%	Pred. No.	3.4e+08						
Matches	204	Conservative	70	Mismatches	95	Indels	8	Gaps	

1 MG1KGLTKLADNAPKAMKFOKFSSYFGKIAVVASMSIYOFI.IVVCRTGMEILINEACE

```

Db 1 MGTHGLAKLLADVAPSAIRENDIKSYFGKRVKALDAASMSYQFLIAT-KGGGVYQNEBGF 59
Qy 61 VTSHLQGMFNRIITRLLEAGIKPVYVFDGKRPDOKKQELAKRYSKRDCACTKDLTEAVEVGD 120
Db 60 TTSHLMGMFYRTIRKMGNGVYKPVYVFDGKRPOLKSGELAKRSERBRAEAKKQIQQAQADAG 119
Qy 121 KDAIEKLSKRTVAVTQHNEDCKRLRLKMGVYVVEAFSAFACCAALCNDKVFVAASGD 180
Db 120 EEEVEKTKRLVAVTKQHNEDCKRLRLKMGVYVVEAFSAFACCAALCNDKVFVAASGD 179
Qy 181 MDLTFGAPRFLRLHMDPSSSKIPVMEFDVAKVLELELTMDQFIDICTLGGCCYCSIK 240
Db 180 MDCLTFGSPVLMRHLTASFAKKLPIQEFHLISVLOELGMDQFVLCILGSDYCFESIR 239
Qy 241 GIGGQIALKLIROHGSTLELLENKNDROYIPEDMPYQEARLIFKEPNVTLD--IPELKW 299
Db 240 GIGKRAVADLTQKHKSIEETVRLDPSKYPVPEMMLHKEAQOLFLEPEV-LDPESVELKW 299
Qy 299 TAPDEGLISFLVNDGNEFDVTKATEIKISAKNKSOGRLSEFPKPTATSAFLKKE 358
Db 299 SEPEELVAFMGCEKQFSEERIRSGVKRLSKRSOGSTQGRLDLDFKVTGSLSS-AKRKE 357
Qy 359 TSDKTSKAANKTKAGGKK 379
Db 358 PE--PKGPARRKAKTGAGCK 375

RESULT 10
QyJHW7 PRELIMINARY: PRI: 380 AA.
AC QyJHW7
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-OCT-2002 (TREMblrel. 22, Last annotation update)
DE Flag structure-specific endonuclease.
CN FEN-1
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Chen D., Cao G., Yang S., Li M., Chen J.;
RT "Cloning and characterization of a rat DNA structure-specific
RT endonuclease (Fen-1).";
KL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF201018; AAF01265.1;
DR HSSP: Q58839.1A76;
DR InterPro: IPR002421; 5.3_exonuclease;
DR InterPro: IPR000513; Fco_N_Lf;
DR InterPro: IPR003584; HNH_2;
DR InterPro: IPR006084; XPG_Rad;
DR InterPro: IPR006085; XPG_L;
DR Pfam: PF01367; 5.3_exonuclease;
DR Pfam: PF00867; XPG_I;
DR Pfam: PF00752; XPG_N;
DR PRINTS: PR00853; XTRADSUPER;
DR SMART: SM00279; HNH2;
DR SMART: SK00484; XPG1;
DR SMART: SK00485; XPGN;
DR PROSITE: PS00842; XPG_2;
KM Endonuclease.
SQ
SEQUENCE 380 AA: 42622 MW: 61698CET16F182136 CRC64;

```

Query Match 53.5%; Score 1040; DB 11; Length 380;  
 Best Local Similarity 52.6%; Pred. No. 4.7e-68;  
 Matches 200; Conservative 73; Mismatches 101; Indels 6; Gaps 4;

```

Qy 1 MGIGLTKLLADNPKAMKEQFESYFGKRIAVDASSTYQFLIVGRIGTETLTNEAGE 60
Db 1 MGIGLAKLLADVAPSAIRENDIKSYFGKRVKALDAASMSYQFLIAT-KGGGVYQNEBGF 59

```

```

Qy 61 VTSHLQGMFNRIITRLLEAGIKPVYVFDGKRPDOKKQELAKRYSKRDCACTKDLTEAVEVGD 120
Db 60 TTSHLMGMFYRTIRKMGNGVYKPVYVFDGKRPOLKSGELAKRSERBRAEAKKQIQQAQADAG 119
Qy 121 KDAIEKLSKRTVAVTQHNEDCKRLRLKMGVYVVEAFSAFACCAALCNDKVFVAASGD 180
Db 120 EEEVEKTKRLVAVTKQHNEDCKRLRLKMGVYVVEAFSAFACCAALCNDKVFVAASGD 179
Qy 181 MDLTFGAPRFLRLHMDPSSSKIPVMEFDVAKVLELELTMDQFIDICTLGGCCYCSIK 240
Db 180 MDCLTFGSPVLMRHLTASFAKKLPIQEFHLISVLOELGMDQFVLCILGSDYCFESIR 239
Qy 241 GIGGQIALKLIROHGSTLELLENKNDROYIPEDMPYQEARLIFKEPNVTLD--IPELKW 299
Db 240 GIGKRAVADLTQKHKSIEETVRLDPSKYPVPEMMLHKEAQOLFLEPEV-LDPESVELKW 299
Qy 299 TAPDEGLISFLVNDGNEFDVTKATEIKISAKNKSOGRLSEFPKPTATSAFLKKE 358
Db 299 SEPEELVAFMGCEKQFSEERIRSGVKRLSKRSOGSTQGRLDLDFKVTGSLSS-AKRKE 357
Qy 359 TSDKTSKAANKTKAGGKK 379
Db 358 PE--PKGPARRKAKTGAGCK 375

RESULT 11
QyC952 PRELIMINARY: PRI: 380 AA.
AC QyC952
DT 01-MAR-2003 (TREMblrel. 23, Created)
DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE Flag structure specific endonuclease 1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRATIN-C57BL/6J; Tissue=Cerebellum;
RC MEDLINE-22354083; PubMed-12465851;
RA "The FANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
KL Nature 420:565-573(2002);
LR EMBL: AK042942; F003145.2.1;
SQ
SEQUENCE 380 AA: 42589 MW: 63488B271885686E CRC64;

```

Query Match 53.5%; Score 1037; DB 11; Length 380;  
 Best Local Similarity 52.6%; Pred. No. 7.8e-68;  
 Matches 204; Conservative 69; Mismatches 100; Indels 8; Gaps 5;

```

Qy 1 MGIGLTKLLADNPKAMKEQFESYFGKRIAVDASSTYQFLIVGRIGTETLTNEAGE 60
Db 1 MGTHGLAKLLADVAPSAIRENDIKSYFGKRVKALDAASMSYQFLIAT-KGGGVYQNEBGF 59
Qy 61 VTSHLQGMFNRIITRLLEAGIKPVYVFDGKRPDOKKQELAKRYSKRDCACTKDLTEAVEVGD 120
Db 60 TTSHLMGMFYRTIRKMGNGVYKPVYVFDGKRPOLKSGELAKRSERBRAEAKKQIQQAQADAG 119
Qy 121 KDAIEKLSKRTVAVTQHNEDCKRLRLKMGVYVVEAFSAFACCAALCNDKVFVAASGD 180
Db 120 EEEVEKTKRLVAVTKQHNEDCKRLRLKMGVYVVEAFSAFACCAALCNDKVFVAASGD 179
Qy 181 MDLTFGAPRFLRLHMDPSSSKIPVMEFDVAKVLELELTMDQFIDICTLGGCCYCSIK 240
Db 180 MDCLTFGSPVLMRHLTASFAKKLPIQEFHLISVLOELGMDQFVLCILGSDYCFESIR 239
Qy 241 GIGGQIALKLIROHGSTLELLENKNDROYIPEDMPYQEARLIFKEPNVTLD--IPELKW 299
Db 240 GIGKRAVADLTQKHKSIEETVRLDPSKYPVPEMMLHKEAQOLFLEPEV-LDPESVELKW 299

```

QY 259 TAPDEGLISFLVKNQGNFEDRVTKATEIKSAKKNSSQGLSEFEKPTATISAPLKKRK 356  
 DB 299 SEPNEELFKVMCGKQFSERIRSGVKRLSKSGSTQGLDQFPKVTGSLSS-ARRKE 357  
 QY 359 TSDKTSKAAANKKTKAGSKK 379  
 DB 358 PE---PKCPAKKKAKTGAGK 375

## RESULT 12

ID Q9N3T2 PRELIMINARY: PRT: 382 AA.

AC Q9N3T2  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Hypothetical 42.5 kDa protein.  
 GN Y4766A.8.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematozoa; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Pelodermatidae; Caenorhabditis.  
 CX NCBI:TaxID:6239;

RN 11)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=930613; PubMed=9851916;  
 RA Note:

RT "Genome sequence of the nematode C. elegans: a platform for  
 RT investigating biology. The C. elegans Sequencing Consortium."  
 RT Science 282:2012-2018(1998).  
 RN 12)

RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RT "The sequence of C. elegans cosmid Y4766A."  
 RT Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
 RN 13)

RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Waterston R.;  
 RT "Direct Submission."  
 RT Submitted (AUG-2001) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AC024751; AAF60553.1;

DR HSSP: Q58839; 1A76.  
 DR Wormpep: Y4766A.8; CP22109.

DR InterPro: IPR002421; 5\_3\_exonuclease.

DR InterPro: IPR000513; Exo\_N\_1.

DR InterPro: IPR003584; HNH\_2.

DR InterPro: IPR006084; XPG\_Rad.

DR InterPro: IPR006085; XPG\_I.

DR InterPro: IPR006085; XPG\_N.

DR Pfam: PF01367; 5\_3\_exonuclease; 1.

DR Pfam: PF00867; XPG\_1; 1.

DR Pfam: PF00752; XPG\_N\_1; 1.

DR PRINTS: PR00853; XPGRADSUPPER.

DR SMART: SM00279; HNH2; 1.

DR SMART: SM00484; XPGI; 1.

DR SMART: SM00485; XPGN; 1.

DR SMART: SM00485; XPGN; 1.

DR SMART: SM00485; XPGN; 1.

DR SMART: SM00485; XPGN; 1.

DR SMART: SM00485; XPGN; 1.

DR SMART: SM00485; XPGN; 1.

DR SMART: SM00485; XPGN; 1.

DR SMART: SM00485; XPGN; 1.

DR SMART: SM00485; XPGN; 1.

DR SMART: SM00485; XPGN; 1.

DR SMART: SM00485; XPGN; 1.

DR SMART: SM00485; XPGN; 1.

DR SMART: SM00485; XPGN; 1.

DR SMART: SM00485; XPGN; 1.

DR SMART: SM00485; XPGN; 1.

DR SMART: SM00485; XPGN; 1.

DR SMART: SM00485; XPGN; 1.

DR SMART: SM00485; XPGN; 1.

DR SMART: SM00485; XPGN; 1.

QY 121 KDAFKLSKRTKVTROENBQKFLR:MGVYVAPASFAEPAACIINDKPAVASPD  
 DB 120 VKEAKFERKRVKTKQONPAKRLGIMGIVVEAPENAAQAHVYKAGKFGIVTED  
 QY 181 MSLTFCAPRPRLHMDPSSKRIIVMEFDVAKVLEETFLTMDQFDICIGQDYSIK  
 DB 180 MDALTFSTVLRLHFLAVAKKIKIKERNISIAIEBKISVEEITDCLILGQYCTIR

## RESULT 13

ID Q9N3T2 PRELIMINARY: PRT: 378 AA.

AC Q9N3T2  
 DT 01-OCT-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Hyp endonuclease-1.  
 GN FEN1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CX NCBI:TaxID:10090;

RN 11)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129sv/c;  
 RA KaranJawala Z.E.; Shi X.; Hsieh C.-T.; Hsieh M.K.;  
 RT "The Mammalian FEN1 Locus: Structure and Conserved Sequence  
 RT Features."  
 RT Microb. Comp. Genomics 0:0-0(2001).  
 DR EMBL: AY014962; AAK01853.1;

DR HSSP: Q58839; 1A76.  
 DR InterPro: IPR002421; 5\_3\_exonuclease.

DR InterPro: IPR000513; Exo\_N\_1.

DR InterPro: IPR003584; HNH\_2.

DR InterPro: IPR006084; XPG\_Rad.

DR InterPro: IPR006085; XPG\_I.

DR InterPro: IPR006085; XPG\_N.

DR Pfam: PF01367; 5\_3\_exonuclease; 1.

DR Pfam: PF00867; XPG\_1; 1.

DR Pfam: PF00752; XPG\_N\_1; 1.

DR PRINTS: PR00853; XPGRADSUPPER.

DR SMART: SM00279; HNH2; 1.

DR SMART: SM00464; XPGI; 1.

DR SMART: SM00485; XPGN; 1.

DR PROSITE: PS00341; XPG\_1; 1.

DR PROSITE: PS00341; XPG\_1; 1.

DR PROSITE: PS00341; XPG\_1; 1.

DR PROSITE: PS00341; XPG\_1; 1.

DR PROSITE: PS00341; XPG\_1; 1.

DR PROSITE: PS00341; XPG\_1; 1.

DR PROSITE: PS00341; XPG\_1; 1.

DR PROSITE: PS00341; XPG\_1; 1.

DR PROSITE: PS00341; XPG\_1; 1.

DR PROSITE: PS00341; XPG\_1; 1.

DR PROSITE: PS00341; XPG\_1; 1.

DR PROSITE: PS00341; XPG\_1; 1.

DR PROSITE: PS00341; XPG\_1; 1.

DR PROSITE: PS00341; XPG\_1; 1.

DR PROSITE: PS00341; XPG\_1; 1.

DR PROSITE: PS00341; XPG\_1; 1.

DR PROSITE: PS00341; XPG\_1; 1.

DR PROSITE: PS00341; XPG\_1; 1.

DR PROSITE: PS00341; XPG\_1; 1.

DR PROSITE: PS00341; XPG\_1; 1.

DR PROSITE: PS00341; XPG\_1; 1.

DR PROSITE: PS00341; XPG\_1; 1.

DR PROSITE: PS00341; XPG\_1; 1.

DR PROSITE: PS00341; XPG\_1; 1.

DR PROSITE: PS00341; XPG\_1; 1.

DR PROSITE: PS00341; XPG\_1; 1.

DR PROSITE: PS00341; XPG\_1; 1.

DR PROSITE: PS00341; XPG\_1; 1.

DR PROSITE: PS00341; XPG\_1; 1.

DR PROSITE: PS00341; XPG\_1; 1.

DR PROSITE: PS00341; XPG\_1; 1.

DR PROSITE: PS00341; XPG\_1; 1.

DR PROSITE: PS00341; XPG\_1; 1.

```

QY 121 KDAIEKLSKRTVAVTROHNEDECKRLILKMGVPVFAVPAFAACALCINQVFAVASED 180
DB 118 EEEVFKETKRLVAVTQHNDECHLSTKMGIPRLAPSAFASCAALACAGVAAATD 177
QY 181 MSLLTGAPFRLHLMDPSSKRIPVMEFVAVLELELTMDQFIDLTLCCGDCYDSIK 240
DB 178 MCLTFSPVLMHLTASEAKKLPTEGFHLSRYLTQELGNQFQFDCLLSDPYCESIA 237
QY 241 GIGGCTAKLIRHGSFESTLENNKDRQVPEDMVYQARRLFKFEPNVT-LDIFLKAT 259
DB 238 GIGAKRAVDLIQHKSLIEELVRDLPSKIPVPENMHLKDAQJLFEFVQPESEVELKWS 297
QY 300 APDEEGLISFLVNDGNFEDRVKALEIKISAKNKSQGLESPFKPATISAPLSKRT 359
DB 298 EPNDEELVKFMCCEKQFSEERISISGVKRLSKNSQSTQGRUDLPFNVCSSLS-AKRFKP 356
QY 360 SDKTSKAAANKTKKAGKK 379
DB 357 E---PKGPARKKAKTGAGK 373

RESULT 14
C96830
ID 096830 PRELIMINARY: PRT: 385 AA.
AC 096830:
DT 01-MAY-1999 (TEMBLrel: 10, Created)
DT 01-MAY-1999 (TEMBLrel: 10, Last sequence update)
DT 01-OCT-2002 (TEMBLrel: 22, Last annotation update)
DE EG:EG0003.3 protein.
GN FEN1 OR EG:EG0003.3 OR CG8648.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Spizididae; Drosophilidae; Drosophila.
OX NCBI_TaxID:7227;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY:
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gacygne J.D.,
RA Anagnostides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazes R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter P.G., Holt G., Nelson C.R., Miles G.L.C.,
RA April J.F., Agbayani A., An H.-J., Andrews Pannkoeh C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale C., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borczyk D., Botchan M.R., Bouck J., Brooksstein P., Broshkov S.,
RA Britis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra S.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Schoha S., Dankov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischman W.,
RA Fodler C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Gloor A., Gong F., Gorell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman J.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howard T.J., Wei M.-H., Ibeagwa C.,
RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Metuliev G., Mishina N.V., Mobarly C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.J.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M.C., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

```

```

RA Ye J., Yeh R.-F., Zaveri J.S., Zhu M., Zhao S., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.X., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.K., Venter J.C.,
RA "The genome sequence of Drosophila melanogaster."
RA Science 287:2185-2195(2000).
RN 121
RP SEQUENCE FROM N.A.
RA Murphy L., Harris D., Bartell B.,
RA "Sequencing the distal X chromosome of Drosophila melanogaster."
RA Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN 131
RP SEQUENCE FROM N.A.
RA Bios P.,
RA Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DB EMBL: A603805; AF57544.1;
DB EMBL: A603863; CA21120.1;
DB HSPB: Q58839.1A76;
DB FLYbase: FBgn0025852; FEN1;
DB Interpro: IPR002421; 5_3_exonuclease;
DB Interpro: IPR00513; EXO_N1;
DB Interpro: IPR003583; HHH_1;
DB Interpro: IPR003584; HHH_2;
DB Interpro: IPR006084; XPGC_Nad;
DB Interpro: IPR006085; XPG_L;
DB Interpro: IPR006085; XPG_N;
DB Pfam: PF01367; 5_3_exonuclease; 1;
DB Pfam: PF00867; XPG_L;
DB Pfam: PF00752; XPG_N;
DB PRINTS: PR00853; XPGRADSUPER;
DB SMART: SM00278; HNH1; 1;
DB SMART: SM00279; HNH2; 1;
DB SMART: SM00484; XPG1; 1;
DB SMART: SM00485; XPGN; 1;
DB PROSITE: PS00842; XPG_1; 1;
DB PROSITE: PS00842; XPG_2; 1;
DB SEQUENCE 385 AA; 42548 MW; DDA0831C2BDA9240 CRC64;

Query Match 51.5%; Score 958.5; DR 5; Length 385;
Best Local Similarity 52.9%; Freq. No. 5,26,65;
Matches 202; Conservative 65; Mismatches 104; Indels 11; Gaps

QY 1 MGIGKSLKLIADNAPVAPKAPKPFESYFCKKIVADASKSYQPLIVAGTGMETTCNFAE 6
DB 1 MGIGKSLKLIADNAPVAPKAPKPFESYFCKKIVADASKSYQPLIVAGTGMETTCNFAE 6
QY 61 VTSKQGMENKTRIL:EAQIKPVPVYVAKAPRPMKKEKLAKRSKRDA-KLLEAVFVCD 1
DB 61 VTSKQGMENKTRIL:EAQIKPVPVYVAKAPRPMKKEKLAKRSKRDA-KLLEAVFVCD 1
DB 60 PISHLKMGFRTIRLLNGIKFPYVVDGKPPDKSGELAKFAFREAEKALKAATDAGD 1
QY 121 KDAIEKLSKRTVAVTROHNEDECKRLILKMGVPVFAVPAFAACALCINQVFAVASED 1
DB 121 KDAIEKLSKRTVAVTROHNEDECKRLILKMGVPVFAVPAFAACALCINQVFAVASED 1
QY 181 MSLLTGAPFRLHLMDPSSKRIPVMEFVAVLELELTMDQFIDLTLCCGDCYDSIK 2
DB 181 MSLLTGAPFRLHLMDPSSKRIPVMEFVAVLELELTMDQFIDLTLCCGDCYDSIK 2
QY 241 GIGGCTAKLIRHGSFESTLENNKDRQVPEDMVYQARRLFKFEPNVT-LDIFLKAT 2
DB 241 GIGGCTAKLIRHGSFESTLENNKDRQVPEDMVYQARRLFKFEPNVT-LDIFLKAT 2
QY 300 APDEEGLISFLVNDGNFEDRVKALEIKISAKNKSQGLESPFKPATISAPLSKRT 3
DB 300 APDEEGLISFLVNDGNFEDRVKALEIKISAKNKSQGLESPFKPATISAPLSKRT 3
QY 360 SDKTSKAAANKTKKAGKK 376
DB 360 SDKTSKAAANKTKKAGKK 376
QY 357 E---PKGPARKKAKTGAGK 373
DB 357 E---PKGPARKKAKTGAGK 373

RESULT 15
C96830
ID C96830 PRELIMINARY: PRT: 650 AA.

```

AC Q96Z01;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Flap endonuclease-1.  
GN FEN1.  
OS Plasmodium falciparum.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_Taxid=5833;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Casta J.J. III, Schmutte C., Taraschi T.F.;  
RT \*Flap Endonuclease-1 (Plasmodium falciparum).  
RL Submitted (JUN-2000) to the EMBL/GenBank/EBI databases.  
DR EMBL: AF278764; AAC01445.1; .  
DR HSSP: Q58839; A76.  
DR InterPro: IPR002421; 5\_3\_exonuclease.  
DR InterPro: IPR000513; EXO\_N\_F.  
DR InterPro: IPR003584; HNH\_2.  
DR InterPro: IPR006084; XPG\_Kad.  
DR InterPro: IPR006086; XPG\_I.  
DR InterPro: IPR006085; XPG\_N.  
DR Pfam: PF01567; 5\_3\_exonuclease; 1.  
DR Pfam: PF00867; XPG\_1; .  
DR Pfam: PF00752; XPG\_N; 1.  
DR PRINTS: PR00853; XPGRADSUPER.  
DR SMART: SM00279; HNH2; .  
DR SMART: SM00484; XPG1; .  
DR SMART: SM00485; XPGN; 1.  
DR Endonuclease.  
KW Endonuclease.  
SQ SEQUENCE 650 AA: 73989 MW: 2P9464DAB7F45B1D CRC64;  
  
Query Match 45.7%; Score 887; DB: 5; Length 650;  
Best Local Similarity 46.6%; Pred. No. 1,5e-56;  
Matches 164; Conservative 74; Mismatches 117; Indels 20; Gaps 5;  
  
QY 1 MGKIGTLKLADAPNAKMEQKFESEYFGKRIADVADSMSTYQFLIVGRGM-EILTNEAG 59  
DB 1 MGKIGTLKLADAPNAKMEQKFESEYFGKRIADVADSMSTYQFLIVGRGM-EILTNEAG 59  
QY 60 EYTSHQGFNRTIRLENGIKPVYVFCGKPPDMKKQELAKRYSKRDATKDL-EAVVEG 119  
DB 61 EYTSHQGFNRTIRLENGIKPVYVFCGKPPDMKKQELAKRYSKRDATKDL-EAVVEG 119  
QY 120 DKDALEKLSKRTVYKTRQHNEDCKRLRLMGVYVVEAPSEAECAALCINKEVFAVASE 175  
DB 121 NLEIKKQSGRTVYKTRKQNEEAKKLLTLMGIPITPAPEAESOCAPLIRYLAHATATE 180  
QY 180 DMDSLTFGAPFLRLHL-----MDPSKKIPVMEFDYAKVLEELTMDQFIDLC 228  
DB 181 DADALVEFGTKILIRNLNANATSNQNNKNSKRGYILTEINDEYIKGILNIMDEFIDC 240  
QY 229 ILGGCDYCSIKIGGOTALK--IRQGSIESILENLNDRQCIPEDMPYQEARRLFKEPN 288  
DB 241 ILGGCDYCSIKIGGOTALK--IRQGSIESILENLNDRQCIPEDMPYQEARRLFKEPN 288  
QY 289 VLDLP-----ELKWTAPDEGLISFLVDCNGFNEDRVTKALEIKSANKKSSQGLESEF 344  
DB 301 V---LPKEDIKIDWNEPQIEELKHFLLIKQYNENELRVNTYINRLKARKVTCRRIDNPF 357  
QY 345 KPIATTSAPLARKKT--SDKTSKAANKKTIKAGGK 378  
DB 358 TACTKSKTKLIVEITKEQTLPAKGGKRRPTAGDK 392

Search completed: October 22, 2003, 12:42:37  
Job time : 104 secs



Db 361 DKTSKAANKTKAGCKK 379

# RESULT 2

US-09-426-557-6  
 : Sequence 6, Application US/09426557  
 : Patent No. 6232527  
 : GENERAL INFORMATION:  
 : APPLICANT: Mahajan, Pramod B.  
 : TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses  
 : TITLE OF INVENTION: thereof  
 : FILE REFERENCE: 0961  
 : CURRENT APPLICATION NUMBER: US/09/426,557  
 : CURRENT FILING DATE: 1999-10-22  
 : EARLIER APPLICATION NUMBER: 60/112,332  
 : EARLIER FILING DATE: 1998-12-15  
 : NUMBER OF SEQ ID NOS: 10  
 : SOFTWARE: FastSeq for Windows Version 3.0  
 : SEQ ID NO: 6  
 : LENGTH: 379  
 : TYPE: PRI  
 : ORGANISM: Zea mays  
 : US-09-426-557-6

Query Match: 100.0%; Score 1939; DB 3; Length 379;  
 Best Local Similarity 100.0%; Pred. No. 1,9e-183;  
 Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGKGLTKLADNAPKAKKEQFESEYFGKRIADVDSMSIYQFLIVGRGEMETLNEAGE 60  
 DB 1 MGKGLTKLADNAPKAKKEQFESEYFGKRIADVDSMSIYQFLIVGRGEMETLNEAGE 60  
 QY 61 VTSHLOGMERNRTIRLEAGIKIPYVVDGKPPDMKKOELAKRYSKRDATKDLTEAVEVG 123  
 DB 61 VTSHLOGMERNRTIRLEAGIKIPYVVDGKPPDMKKOELAKRYSKRDATKDLTEAVEVG 123  
 QY 121 KDAIEKLSKRTVAVTQHNEDECKRLRLMGVYVVAEPAEACALCINDKVFVAASD 189  
 DB 121 KDAIEKLSKRTVAVTQHNEDECKRLRLMGVYVVAEPAEACALCINDKVFVAASD 189  
 QY 181 MDLTFGAPREFLRHMDPSSKRIYVMEFDVAKYLELELTMDOFIDLCICGDCYDSIK 240  
 DB 181 MDLTFGAPREFLRHMDPSSKRIYVMEFDVAKYLELELTMDOFIDLCICGDCYDSIK 240  
 QY 241 GIGGOTALKLRHOGSISTESTLENKNDROYIPEDMWYQEARRLFKEPNVTLIDIPELKMTA 300  
 DB 241 GIGGOTALKLRHOGSISTESTLENKNDROYIPEDMWYQEARRLFKEPNVTLIDIPELKMTA 300  
 QY 301 PDEGLISFLVKGNGFNEEDRVTALEKIKSAKSSQGRLESFFKPTATTSAPLKRKETS 360  
 DB 301 PDEGLISFLVKGNGFNEEDRVTALEKIKSAKSSQGRLESFFKPTATTSAPLKRKETS 360  
 QY 361 DKTSKAANKTKAGCKK 379  
 DB 361 DKTSKAANKTKAGCKK 379

RESULT 3  
 : US-09-426-557-4  
 : Sequence 4, Application US/09426557  
 : Patent No. 6232527  
 : GENERAL INFORMATION:  
 : APPLICANT: Mahajan, Pramod B.  
 : TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses  
 : TITLE OF INVENTION: thereof  
 : FILE REFERENCE: 0961  
 : CURRENT APPLICATION NUMBER: US/09/426,557  
 : CURRENT FILING DATE: 1999-10-22  
 : EARLIER APPLICATION NUMBER: 60/112,332  
 : EARLIER FILING DATE: 1998-12-15  
 : NUMBER OF SEQ ID NOS: 10  
 : SOFTWARE: FastSeq for Windows Version 3.0

: SEQ ID NO: 4  
 : LENGTH: 379  
 : TYPE: PRI  
 : ORGANISM: Zea mays  
 : US-09-426-557-4

Query Match: 99.7%; Score 1933; DB 3; Length 379;  
 Best Local Similarity 99.7%; Pred. No. 7.6e-183;  
 Matches 379; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGKGLTKLADNAPKAKKEQFESEYFGKRIADVDSMSIYQFLIVGRGEMETLNEAGE 60  
 DB 1 MGKGLTKLADNAPKAKKEQFESEYFGKRIADVDSMSIYQFLIVGRGEMETLNEAGE 60  
 QY 61 VTSHLOGMERNRTIRLEAGIKIPYVVDGKPPDMKKOELAKRYSKRDATKDLTEAVEVG 123  
 DB 61 VTSHLOGMERNRTIRLEAGIKIPYVVDGKPPDMKKOELAKRYSKRDATKDLTEAVEVG 123  
 QY 121 KDAIEKLSKRTVAVTQHNEDECKRLRLMGVYVVAEPAEACALCINDKVFVAASD 189  
 DB 121 KDAIEKLSKRTVAVTQHNEDECKRLRLMGVYVVAEPAEACALCINDKVFVAASD 189  
 QY 181 MDLTFGAPREFLRHMDPSSKRIYVMEFDVAKYLELELTMDOFIDLCICGDCYDSIK 240  
 DB 181 MDLTFGAPREFLRHMDPSSKRIYVMEFDVAKYLELELTMDOFIDLCICGDCYDSIK 240  
 QY 241 GIGGOTALKLRHOGSISTESTLENKNDROYIPEDMWYQEARRLFKEPNVTLIDIPELKMTA 300  
 DB 241 GIGGOTALKLRHOGSISTESTLENKNDROYIPEDMWYQEARRLFKEPNVTLIDIPELKMTA 300  
 QY 301 PDEGLISFLVKGNGFNEEDRVTALEKIKSAKSSQGRLESFFKPTATTSAPLKRKETS 360  
 DB 301 PDEGLISFLVKGNGFNEEDRVTALEKIKSAKSSQGRLESFFKPTATTSAPLKRKETS 360  
 QY 361 DKTSKAANKTKAGCKK 379  
 DB 361 DKTSKAANKTKAGCKK 379

RESULT 4  
 : US-09-426-557-6  
 : Sequence 8, Application US/09426557  
 : Patent No. 6232527  
 : GENERAL INFORMATION:  
 : APPLICANT: Mahajan, Pramod B.  
 : TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses  
 : TITLE OF INVENTION: thereof  
 : FILE REFERENCE: 0961  
 : CURRENT APPLICATION NUMBER: US/09/426,557  
 : CURRENT FILING DATE: 1999-10-22  
 : EARLIER APPLICATION NUMBER: 60/112,332  
 : EARLIER FILING DATE: 1998-12-15  
 : NUMBER OF SEQ ID NOS: 10  
 : SOFTWARE: FastSeq for Windows Version 3.0  
 : SEQ ID NO: 8  
 : LENGTH: 379  
 : TYPE: PRI  
 : ORGANISM: Zea mays  
 : US-09-426-557-6

Query Match: 99.7%; Score 1933; DB 3; Length 379;  
 Best Local Similarity 99.7%; Pred. No. 7.6e-183;  
 Matches 379; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGKGLTKLADNAPKAKKEQFESEYFGKRIADVDSMSIYQFLIVGRGEMETLNEAGE 60  
 DB 1 MGKGLTKLADNAPKAKKEQFESEYFGKRIADVDSMSIYQFLIVGRGEMETLNEAGE 60  
 QY 61 VTSHLOGMERNRTIRLEAGIKIPYVVDGKPPDMKKOELAKRYSKRDATKDLTEAVEVG 123  
 DB 61 VTSHLOGMERNRTIRLEAGIKIPYVVDGKPPDMKKOELAKRYSKRDATKDLTEAVEVG 123  
 QY 121 KDAIEKLSKRTVAVTQHNEDECKRLRLMGVYVVAEPAEACALCINDKVFVAASD 189  
 DB 121 KDAIEKLSKRTVAVTQHNEDECKRLRLMGVYVVAEPAEACALCINDKVFVAASD 189



```

Db      221 KDAIEKLSKRTVKTROHNEDECKRLRLMGVYVVEAPSEAECAALCINRQVAVASD 180
Cy      181 MDSLTFGARFRLRLHMDPSSKRIPIWEPDVAKVLELELTDQFDLCILGCDYCSIK 240
Db      181 KDSLTFGARFRLRLHMDPSSKRIPIWEPDVAKVLELELTDQFDLCILGCDYCSIK 240
Cy      241 GIGGOTALKLIRHOSISILNENKDRQYQPEDKPYQDARLEFEKPVNTLIDIFLAKIA 300
Db      241 GIGGOTALKLIRHOSISILNENKDRQYQPEDKPYQDARLEFEKPVNTLIDIFLAKIA 300
Cy      301 PDEGLISFVKNQNFNEEDRVKALEIKISAKNKSOGRLSEFFKPTATVSNAPJRKRETS 360
Db      301 PDEGLISFVKNQNFNEEDRVKALEIKISAKNKSOGRLSEFFKPTATVSNAPJRKRETS 360
Cy      361 DKTSKAANKTKKAGKKK 379
Db      361 DKTSKAANKTKKAGKKK 379

```

# RESULT 5

```

US-08-823-516-137
Sequence 137, Application US/08823516
Patent No. 5994069
GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
APPLICANT: Lyamichev, Victor I.
APPLICANT: Most, Andrea L.
APPLICANT: Brown, Maty Ann D.
TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medien 6 Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/756,386
FILING DATE: 28-NOV-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/662,853
FILING DATE: 12-JUL-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/559,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Lyolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: F0RS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
FAX: (415) 397-8338

```

INFORMATION FOR SEQ ID NO: 137:

SEQUENCE CHARACTERISTICS:

LENGTH: 380 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: not relevant

MOLECULE TYPE: protein

US-08-823-516-137

Query Match: 53.2%, Score 1032, DB 2, Length 380;

Best Local Similarity: 53.8%, Prod. No. 93;

Matches: 206; Conservative: 69; Mismatches: 98; Indels: 10; Gaps:

```

Cy      1 MSLKGLKILADNAPKAKKQKQKFSYGRKILAVASMSLYQFLVAGTCMETLINAGE
Db      1 MSLKGLKILADNAPKAKKQKQKFSYGRKILAVASMSLYQFLVAGTCMETLINAGE
Cy      61 VSHIQGFNFTIRLENGIKPVYVVDKPPKPKKQKFLAKRYSKRDATKDLTAVEYGD
Db      60 TSHLKEKYWTILMMENGKIFVYVVDKPPKPKKQKFLAKRYSKRDATKDLTAVEYGD
Cy      121 KDAIEKLSKRTVKTROHNEDECKRLRLMGVYVVEAPSEAECAALCINRQVAVASD
Db      120 EDEVEKELTKRLVAVKQHNQDCKHLISMGVYVVEAPSEAECAALCINRQVAVASD
Cy      181 MDSLTFGARFRLRLHMDPSSKRIPIWEPDVAKVLELELTDQFDLCILGCDYCSIK
Db      180 MDSLTFGARFRLRLHMDPSSKRIPIWEPDVAKVLELELTDQFDLCILGCDYCSIK
Cy      241 GIGGOTALKLIRHOSISILNENKDRQYQPEDKPYQDARLEFEKPVNTLIDIFLAKIA
Db      240 GIGGOTALKLIRHOSISILNENKDRQYQPEDKPYQDARLEFEKPVNTLIDIFLAKIA
Cy      299 TAPDEGLISFVKNQNFNEEDRVKALEIKISAKNKSOGRLSEFFKPTATVSNAPJRKRE
Db      299 SEPDEGLISFVKNQNFNEEDRVKALEIKISAKNKSOGRLSEFFKPTATVSNAPJRKRE
Cy      359 TSKAANKTKKAGKKK 379
Db      358 PEPKSTIKKKA---KTKAAGKFK 377

```

# RESULT 6

```

US-09-426-557-10
Sequence 10, Application US/09426557
Patent No. 5232557
GENERAL INFORMATION:
APPLICANT: Bahajan, Pramod B.
TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
FILE REFERENCE: 0961
CURRENT APPLICATION NUMBER: US/09/426,557
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 60/112,332
EARLIER FILING DATE: 1998-12-15
SOFTWARE: FASTSTRO for Windows Version 3.0
SEQ ID NO: 10
LENGTH: 380
TYPE: PRT
ORGANISM: Homo sapiens
US-09-426-557-10

```

Query Match: 53.2%, Score 1032, DB 3, Length 380;

Best Local Similarity: 53.8%, Prod. No. 93;

Matches: 206; Conservative: 69; Mismatches: 98; Indels: 10; Gaps:

```

Cy      1 MSLKGLKILADNAPKAKKQKQKFSYGRKILAVASMSLYQFLVAGTCMETLINAGE
Db      1 MSLKGLKILADNAPKAKKQKQKFSYGRKILAVASMSLYQFLVAGTCMETLINAGE
Cy      61 VSHIQGFNFTIRLENGIKPVYVVDKPPKPKKQKFLAKRYSKRDATKDLTAVEYGD
Db      60 TSHLKEKYWTILMMENGKIFVYVVDKPPKPKKQKFLAKRYSKRDATKDLTAVEYGD

```

[illegible]

```

1      RESULT 7
2      US-08-455-963E-1
3      Sequence 1, Application US/08455968E
4      Patent No. 5874283
5      GENERAL INFORMATION:
6      APPLICANT: Harrington, John L.
7      APPLICANT: Hsieh, Chih-Lin
8      APPLICANT: Lieber, Michael
9      TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
10     NUMBER OF SEQUENCES: 63
11     CORRESPONDENCE ADDRESS:
12     ADDRESSEE: Townsend and Townsend and Crew LLP
13     STREET: Two Embarcadero Center, 9th Floor
14     CITY: San Francisco
15     STATE: California
16     COUNTRY: USA
17     ZIP: 94111-3634
18     COMPUTER READABLE FORM:
19     MEDIUM TYPE: Floppy disk
20     COMPUTER: IBM PC compatible
21     OPERATING SYSTEM: PC-DOS/MS-DOS
22     SOFTWARE: Patent Release #1.0, Version #1.25
23     CURRENT APPLICATION DATA:
24     APPLICATION NUMBER: US/08/455,968E
25     FILING DATE: 30-MAY-1995
26     CLASSIFICATION: 435
27     ATTORNEY/AGENT INFORMATION:
28     NAME: Apple, Randolph T.
29     REGISTRATION NUMBER: 36,429
30     REFERENCE/DOCKET NUMBER: 18995-000100
31     TELECOMMUNICATION INFORMATION:
32     TELEPHONE: 415-576-0200
33     TELEFAX: 415-576-0300
34     INFORMATION FOR SEQ ID NO: 1:
35     SEQUENCE CHARACTERISTICS:
36     LENGTH: 380 amino acids
37     TYPE: amino acid
38     STRANDEDNESS: single
39     TOPOLOGY: linear
40     MOLECULE TYPE: protein
41     US-08-455-963E-1

```

Query Match 53.2% Score 1031 DB 2: Length 385:  
Best Local Similarity 53.8% Pred No. 1.3e-93:  
Matches 206; Conservative 69; Mismatches 99; Indels 10; Caps 6

QY 1 MGIGKLTLLADNAKRAKEKEFESYGRKRAVDASMSYVCFLLVVGRTGKMTLLTNEAF 60  
1 MGIGGGLAKIAVAVASAIREFNDIKSYFERKVALDASMSYVDFLLAV PQSGDAVQNEKE 59

[illegible]

```

1  RESCIT 8
2  US-08-823-516-138
3  Sequence 138, Application US/08821516
4  Patent No. 5994053
5  GENERAL INFORMATION:
6  APPLICANT: Hall, Jeff G.
7  APPLICANT: Lyatchev, Victor I.
8  APPLICANT: Mast, Andrea L.
9  APPLICANT: Brown, Mary Ann D.
10 TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
11 TITLE OF INVENTION: Sequential Invasive Cleavages
12 NUMBER OF SEQUENCES: 163
13 CORRESPONDENCE ADDRESS:
14 ADDRESSEE: Medicon & Carroll, LLP
15 STREET: 220 Montgomery Street, Suite 2200
16 CITY: San Francisco
17 STATE: California
18 COUNTRY: United States Of America
19 ZIP: 94104
20 COMPUTER READABLE FORM:
21 MEDIUM TYPE: Floppy disk
22 COMPUTER: IBM PC compatible
23 OPERATING SYSTEM: PC-DOS/MS-DOS
24 SOFTWARE: PatentIn Release #1.0, Version #1.30
25 CURRENT APPLICATION DATA:
26 APPLICATION NUMBER: US/08/823-516
27 FILING DATE: 24-MAR-1997
28 CLASSIFICATION: 435
29 PRIOR APPLICATION DATA:
30 APPLICATION NUMBER: PCT/US97/01072
31 FILING DATE: 21-JAN-1997
32 PRIOR APPLICATION DATA:
33 APPLICATION NUMBER: US 08/759,038
34 FILING DATE: 02-DEC-1996
35 PRIOR APPLICATION DATA:
36 APPLICATION NUMBER: US 08/756,364
37 FILING DATE: 29-NOV-1996
38 PRIOR APPLICATION DATA:
39 APPLICATION NUMBER: US 08/682,854
40 FILING DATE: 12-JUL-1996
41 PRIOR APPLICATION DATA:
42 APPLICATION NUMBER: US 08/599,497
43 FILING DATE: 24-JAN-1995
44 ATTORNEY/AGENT INFORMATION:
45 NAME: Ingolia, Diane E.

```

REGISTRATION NUMBER: 40,027  
 REFERENCE/DOCKET NUMBER: FORS-02736  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 705-8410  
 TELEFAX: (415) 397-8338  
 INFORMATION FOR SEQ ID NO: 138:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 378 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: not relevant  
 TOPOLOGY: not relevant  
 MOLECULE TYPE: protein  
 US-09-823-516-138

Query Match 52.38: Score 1014; DB 2: Length 378;  
 Best Local Similarity 53.24; Pred. No. 6, 1e-92;

Matches 202: Conservative 69; Mismatches 101; Indels 5; Gaps 6;

OY 1 MGIRGLTKLLADNAPKAMKQKESYFGPKRIADASMSIYQFLIVGRTGMEITNEAGE 60  
 DB 1 MGIRGLTKLLADNAPKAMKQKESYFGPKRIADASMSIYQFLIVGRTGMEITNEAGE 59  
 OY 61 VISHLOGMFNRTIRLEAGIKPYVFDGKPPDMKKQELAKRYSKRDATKDTTEAVEVCD 120  
 DB 60 TTS-LMGFPRITR-MENGIRKPYVFDGKPPDMKKQELAKRYSKRDATKDTTEAVEVCD 117  
 OY 121 KDAIEKSKRCVAVTRQHNEDCKRLRLMGVPPVEAPSEAEACALCINDKVFVASED 180  
 DB 118 EEEVEKTKRLVAVTRQHNEDCKRLRLMGVPPVEAPSEAEACALCINDKVFVASED 177  
 OY 181 MDLTFGAPRFLRLMDPSSKRIYVMEFDYAKVLELELMDQFLDCLTLCGQDYDSDIK 240  
 DB 178 MDCLTFGAPRFLRLMDPSSKRIYVMEFDYAKVLELELMDQFLDCLTLCGQDYDSDIK 237  
 OY 241 GIGGOTALKLIRQGSFESTLENLNDROYLPEDMPYQEARRLPKRPNT-IDIPELKMT 299  
 DB 238 GIGAKRAVDLQKHSLEELVRRLDPAKYVPENMLHKEAQOFLLEPFVVDPSVELKMS 297  
 OY 300 APDEGLISFLVKNQNGENRDTKALIKIKSAKNSQGLSEFEPYATTSAPLKRKET 359  
 DB 298 EPNEELVKFMCGKQSEERIRSGVRLSKSGOSTGRLDQFVVTGSLS-AKRRKP 356  
 OY 360 SDKTSKAANKTKAGCKK 379  
 DB 357 E--PKGPARKKAKTGAGK 373

RESULT 9  
 US-08-455-968E-3

Sequence 3, Application US/08455968E  
 Patent No. 5874283

GENERAL INFORMATION:  
 APPLICANT: Harrington, John L.  
 APPLICANT: Hsieh, Chih-Lin  
 APPLICANT: Lieber, Michael  
 TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease  
 NUMBER OF SEQUENCES: 63  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Townsend and Townsend and Crew LLP  
 STREET: Two Embarcadero Center, 8th Floor  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94111-3834

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent's Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/455,968E  
 FILING DATE: 30-MAY-1995  
 CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
 NAME: Apple, Randolph T.  
 REGISTRATION NUMBER: 36,429  
 REFERENCE/DOCKET NUMBER: 18985-000100  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415-576-0200  
 TELEFAX: 415-576-0300  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 377 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-455-968E-3

Query Match 52.08: Score 1007.5; DB 2: Length 377;  
 Best Local Similarity 53.44; Pred. No. 2, 7e-91;

Matches 203: Conservative 68; Mismatches 100; Indels 5; Gaps 6;

OY 1 MGIRGLTKLLADNAPKAMKQKESYFGPKRIADASMSIYQFLIVGRTGMEITNEAGE 60  
 DB 1 MGIRGLTKLLADNAPKAMKQKESYFGPKRIADASMSIYQFLIVGRTGMEITNEAGE 59  
 OY 61 VISHLOGMFNRTIRLEAGIKPYVFDGKPPDMKKQELAKRYSKRDATKDTTEAVEVCD 120  
 DB 60 TTS-LMGFPRITR-MENGIRKPYVFDGKPPDMKKQELAKRYSKRDATKDTTEAVEVCD 117  
 OY 121 KDAIEKSKRCVAVTRQHNEDCKRLRLMGVPPVEAPSEAEACALCINDKVFVASED 180  
 DB 117 MEEVEKTKRLVAVTRQHNEDCKRLRLMGVPPVEAPSEAEACALCINDKVFVASED 177  
 OY 181 MDLTFGAPRFLRLMDPSSKRIYVMEFDYAKVLELELMDQFLDCLTLCGQDYDSDIK 240  
 DB 178 MDCLTFGAPRFLRLMDPSSKRIYVMEFDYAKVLELELMDQFLDCLTLCGQDYDSDIK 237  
 OY 241 GIGGOTALKLIRQGSFESTLENLNDROYLPEDMPYQEARRLPKRPNT-IDIPELKMT 299  
 DB 238 GIGAKRAVDLQKHSLEELVRRLDPAKYVPENMLHKEAQOFLLEPFVVDPSVELKMS 297  
 OY 300 APDEGLISFLVKNQNGENRDTKALIKIKSAKNSQGLSEFEPYATTSAPLKRKET 359  
 DB 298 EPNEELVKFMCGKQSEERIRSGVRLSKSGOSTGRLDQFVVTGSLS-AKRRKP 356  
 OY 360 SDKTSKAANKTKAGCKK 379  
 DB 356 E--PKGPARKKAKTGAGK 372

RESULT 10  
 US-08-455-968E-10

Sequence 10, Application US/08455968E  
 Patent No. 5874283

GENERAL INFORMATION:  
 APPLICANT: Harrington, John L.  
 APPLICANT: Hsieh, Chih-Lin  
 APPLICANT: Lieber, Michael  
 TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease  
 NUMBER OF SEQUENCES: 63  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Townsend and Townsend and Crew LLP  
 STREET: Two Embarcadero Center, 8th Floor  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94111-3834

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent's Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/455,968E

FILING DATE: 30-MAY-1995  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Apple, Randolph T.  
 REGISTRATION NUMBER: 36,429  
 REFERENCE/DOCKET NUMBER: 19985-000100  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415-576-0200  
 TELEFAX: 415-576-0300  
 INFORMATION FOR SEQ ID NO: 10:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 378 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-455-968E-10

Query Match 51.4%; Score 996; DB 2; Length 378;  
 Best Local Similarity 52.4%; Pred. No. 3,7e+90;  
 Matches 199; Conservative 70; Mismatches 103; Indels 8; Gaps 6;

```

QY 1 MGKGLTKLADNAPKAKKEQKESYGRKTAVDASMSYQFLIVGRGTMTLTNAGE 60
DB 1 MEHGIKALADVAPSRRENDIKSYGRKVAIDASMSYQFLIV-RQGADVLQNEGE 59
QY 61 VTSHLQGMENFTIRLENGIKPVYVFDGKPPDMKKQELAKRYSKRDATXDTLEAVEVG 120
DB 60 TTS-LMGMYRTMR-MENGIKPVYVFDGKPPQLKSGELAKRSERRAEKQJQQAQAGM 117
QY 121 KDAIEKLSKRTVKTROHNEDECKRLRLMSVPEVAPSEAECAALCINDKVAVASED 180
DB 118 EEVEKEFTKRLVKTQKHNDECKHLISLMDGIPYLDAPSEASQCALAKKAYAAATFD 177
QY 181 MDLTFGAPRFLRLHMDPSKKIPVMEFDVAKVLEELTMDQFIDLCILCGCYCSK 240
DB 178 MDCLIFGSPVAKRHIASEAKKLQIQRHLSRVLOELDQNEQVDDCILLGSPYCSIR 237
QY 241 GIGGQIALKLRQHSLESLINKRDYQIPEDMFYQEARCKEPPNVT-LDIPELKWT 299
DB 238 GIGAKRAVDLQKHISIEIVRGLDPSKYVPEXWLMHEAQCLPLEBEVYDPEFVFKMS 297
QY 300 APDEGLSFLVKDNGENEDRVTAIEIKSAKAKKSOGRLSEFPKPATSAALKEKE 359
DB 298 EPNEBELVKTMGGEQFEERIRISVVKRLSKRSGSTQGRLDDEFFKVGSLS-AKRREP 356
QY 360 SDKTSKAANKKTKAGCRKK 379
DB 357 F---PKGPAKKKAKTGSAGK 373
  
```

## RESULT 11

US-08-455-968E-5  
 Sequence 5, Application US/08455968E  
 Patent No. 5874283  
 GENERAL INFORMATION:  
 APPLICANT: Harrington, John L.  
 APPLICANT: Hsieh, Chih-Lin  
 APPLICANT: Lieber, Michael  
 TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease  
 NUMBER OF SEQUENCES: 63  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Townsend and Townsend and Crew LLP  
 STREET: Two Embarcadero Center, 8th floor  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94111-3834  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: 09/08/455,968E  
 FILING DATE: 30-MAY-1995  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Apple, Randolph T.  
 REGISTRATION NUMBER: 36,429  
 REFERENCE/DOCKET NUMBER: 19985-000100  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415-576-0200  
 TELEFAX: 415-576-0300  
 INFORMATION FOR SEQ ID NO: 5:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 382 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-455-968E-5

Query Match 48.4%; Score 938; DB 2; Length 382;  
 Best Local Similarity 49.5%; Pred. No. 2e+84;  
 Matches 190; Conservative 65; Mismatches 109; Indels 16; Gaps

```

QY 1 MGKGLTKLADNAPKAKKEQKESYGRKTAVDASMSYQFLIVGRGTMTLTNAGE 60
DB 1 MEHGIKALADVAPSRRENDIKSYGRKVAIDASMSYQFLIV-RQGADVLQNEGE 59
QY 61 VTSHLQGMENFTIRLENGIKPVYVFDGKPPDMKKQELAKRYSKRDATXDTLEAVEVG 120
DB 60 TTS-LMGMYRTMR-MENGIKPVYVFDGKPPQLKSGELAKRSERRAEKQJQQAQAGM 117
QY 121 KDAIEKLSKRTVKTROHNEDECKRLRLMSVPEVAPSEAECAALCINDKVAVASED 180
DB 118 EEVEKEFTKRLVKTQKHNDECKHLISLMDGIPYLDAPSEASQCALAKKAYAAATFD 177
QY 181 MDLTFGAPRFLRLHMDPSKKIPVMEFDVAKVLEELTMDQFIDLCILCGCYCSK 240
DB 178 MDCLIFGSPVAKRHIASEAKKLQIQRHLSRVLOELDQNEQVDDCILLGSPYCSIR 237
QY 241 GIGGQIALKLRQHSLESLINKRDYQIPEDMFYQEARCKEPPNVT-LDIPELKWT 299
DB 238 GIGAKRAVDLQKHISIEIVRGLDPSKYVPEXWLMHEAQCLPLEBEVYDPEFVFKMS 297
QY 300 APDEGLSFLVKDNGENEDRVTAIEIKSAKAKKSOGRLSEFPKPATSAALKEKE 359
DB 298 EPNEBELVKTMGGEQFEERIRISVVKRLSKRSGSTQGRLDDEFFKVGSLS-AKRREP 356
QY 360 SDKTSKAANKKTKAGCRKK 379
DB 357 F---PKGPAKKKAKTGSAGK 373
  
```

## RESULT 12

US-08-823-516-139  
 Sequence 139, Application US/08823516  
 Patent No. 5994869  
 GENERAL INFORMATION:  
 APPLICANT: Hall, Jeff G.  
 APPLICANT: Lyons, Victor J.  
 APPLICANT: Mast, Andrea L.  
 APPLICANT: Bray, Mary Ann G.  
 TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple  
 NUMBER OF SEQUENCES: 163  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Medien & Carroll, LLP  
 STREET: 220 Montgomery Street, Suite 2200  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: United States Of America  
 ZIP: 94104  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/01672
FILING DATE: 21-JAN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane F.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 139:
SEQUENCE CHARACTERISTICS:
LENGTH: 382 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-08-823-516-139

Query Match      48.4% Score 938; DB 2; Length 382;
Best Local Similarity 49.5%; Pred. No. 2e-84;
Matches 190; Conservative 69; Mismatches 109; Indels 16; Gaps 6;

```

```

RESULT 13
US-09-757-653-176
Sequence 176, Application US/08757653
Patent No. 5843669
GENERAL INFORMATION:
APPLICANT: Kaiser, Michael W.
APPLICANT: Lyamichov, Natalia
TITLE OF INVENTION: Cleavage of Nucleic Acid Using
TITLE OF INVENTION: Thermostable Fen-1 Endonucleases
NUMBER OF SEQUENCES: 190
CORRESPONDENCE ADDRESS:
ADDRESS: Medler & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/757,653
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane F.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02565
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 397-8338
TELEFAX: (415) 705-8410
INFORMATION FOR SEQ ID NO: 176:
SEQUENCE CHARACTERISTICS:
LENGTH: 340 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-757-653-176

Query Match      33.4% Score 648.5; DB 2; Length 340;
Best Local Similarity 41.3%; Pred. No. 7.4e-55;
Matches 142; Conservative 62; Mismatches 107; Indels 33; Gaps

```

```

US-08-823-516-79
: Sequence 79, Application US/08823516
: Patent No. 5994069
: GENERAL INFORMATION:
: APPLICANT: Hall, Jeff G.
: APPLICANT: Lyamichev, Victor I.
: APPLICANT: Mast, Andrea L.
: APPLICANT: Brow, Mary Ann D.
: TITLE OF INVENTION: Detection of Nucleic Acids By Multiple
: NUMBER OF SEQUENCES: 163
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Medlen & Carroll, LLP
: STREET: 220 Montgomery Street, Suite 2200
: CITY: San Francisco
: STATE: California
: COUNTRY: United States of America
: ZIP: 94104
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/823,516
: FILING DATE: 24-MAR-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/US97/01072
: FILING DATE: 21-JAN-1997
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/759,038
: FILING DATE: 02-DEC-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/758,314
: FILING DATE: 02-DEC-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/756,386
: FILING DATE: 29-NOV-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/632,853
: FILING DATE: 12-JUL-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/599,491
: FILING DATE: 24-JAN-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Ingolia, Diane E.
: REGISTRATION NUMBER: 40,027
: REFERENCE/DOCKET NUMBER: FORS-02736
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 705-8410
: TELEFAX: (415) 397-8338
: INFORMATION FOR SEQ ID NO: 79:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 340 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-823-516-79

Query Match      33.4%: Score 648.5; DB 2; Length 340;
Best Local Similarity 41.3%; Pred. No. 7.4e-56;
Matches 142; Conservative 62; Mismatches 107; Indels 33; Gaps 6;

```

```

DB      112 1EDAKKLELMGIDIVGAPSCGAAVYAKAGSVASASODYSLLFGAPRLVRLTLT 1
QY      139 SSKKIP-----VMEPVAVLEEEJTKOFTDCTICGCDY-CDSIKGIG 2
DB      132 GKKLPCKNVYVEIKPELTILE---PYLKKLKLIREKLLFALVCTDINPGGIRKIGL 2
QY      245 QVALKLIHQHGSIESITENLRD---RYQJHEDMVYQEARLKEPYNVTDIPELKTAP 3
DB      248 KKALEIVH-----SKDPLAFKQSCVDYVAIKKFFLNPVY-DNYVCVWRDP 2
QY      302 DREGLISTVKNQGNFEDRYTKALEKIKSAKNSQGRLESFFK 345
DB      296 DEFGIKFLCDEHDFSEERVNRGRLKKAIRKSGKSTIFSMFK 339

RESULT 15
US-08-823-516-79
: Sequence 136, Application US/08823516
: Patent No. 5994069
: GENERAL INFORMATION:
: APPLICANT: Hall, Jeff G.
: APPLICANT: Lyamichev, Victor I.
: APPLICANT: Mast, Andrea L.
: APPLICANT: Brow, Mary Ann D.
: TITLE OF INVENTION: Detection of Nucleic Acids By Multiple
: NUMBER OF SEQUENCES: 163
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Medlen & Carroll, LLP
: STREET: 220 Montgomery Street, Suite 2200
: CITY: San Francisco
: STATE: California
: COUNTRY: United States of America
: ZIP: 94104
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/823,516
: FILING DATE: 24-MAR-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/US97/01072
: FILING DATE: 21-JAN-1997
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/759,038
: FILING DATE: 02-DEC-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/758,314
: FILING DATE: 02-DEC-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/756,386
: FILING DATE: 29-NOV-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/632,853
: FILING DATE: 12-JUL-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/599,491
: FILING DATE: 24-JAN-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Ingolia, Diane E.
: REGISTRATION NUMBER: 40,027
: REFERENCE/DOCKET NUMBER: FORS-02736
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 705-8410
: TELEFAX: (415) 397-8338
: INFORMATION FOR SEQ ID NO: 136:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 340 amino acids
: TYPE: amino acid

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## QM protein - protein search, using sw model

Run on: October 22, 2003, 12:40:26 : Search time 31 Seconds  
(without alignments)  
2047.341 Million cell updates/sec

Title: US-09-805-311-2

Perfect score: 1 MKIKGLTKLLADNAPKAKE.....SDKISKAAANKTKAGCKKK 379

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 629382 seqs, 167660630 residues

Total number of hits satisfying chosen parameters: 629382

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

## Database :

Published Applications-AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PC1\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PC105\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1939	100.0	379	9	US-09-805-311-2
2	1939	100.0	379	9	US-09-805-311-6
3	1933	99.7	379	9	US-09-805-311-4
4	1933	99.7	379	9	US-09-805-311-8
5	1032	53.2	380	9	US-09-805-311-10
6	1032	53.2	380	11	US-09-940-244-137
7	1032	53.2	380	11	US-10-290-386-137
8	1032	53.2	380	14	US-10-033-297-137
9	1014	52.3	378	11	US-09-940-244-138
10	1014	52.3	378	12	US-10-290-386-138
11	1014	52.3	378	14	US-10-033-297-138
12	938	48.4	382	11	US-09-940-244-139
13	938	48.4	382	12	US-10-290-386-139
14	938	48.4	382	14	US-10-033-297-139
15	648.5	33.4	340	10	US-09-777-430A-27

16	648.5	33.4	340	11	US-09-940-244-79	Sequence 79, f
17	648.5	33.4	340	11	US-09-940-244-136	Sequence 136, f
18	648.5	33.4	340	12	US-10-290-386-79	Sequence 79, f
19	648.5	33.4	340	12	US-10-290-386-136	Sequence 136, f
20	648.5	33.4	340	14	US-10-033-297-79	Sequence 79, f
21	648.5	33.4	340	14	US-10-033-297-136	Sequence 136, f
22	637	32.9	340	11	US-09-940-244-365	Sequence 365, f
23	632.5	32.6	343	11	US-09-940-244-357	Sequence 357, f
24	626	32.3	340	11	US-09-940-244-384	Sequence 384, f
25	624	32.2	346	11	US-09-940-244-379	Sequence 379, f
26	615	31.7	340	11	US-09-940-244-402	Sequence 402, f
27	615	31.7	351	11	US-09-940-244-337	Sequence 337, f
28	595	30.7	351	11	US-09-940-244-341	Sequence 341, f
29	582	30.0	328	11	US-09-940-244-398	Sequence 398, f
30	572	29.5	350	11	US-09-940-244-394	Sequence 394, f
31	569	29.1	351	11	US-09-940-244-361	Sequence 361, f
32	560.5	28.9	326	10	US-09-777-430A-28	Sequence 28, f
33	560.5	28.9	326	11	US-09-940-244-75	Sequence 75, f
34	560.5	28.9	326	11	US-09-940-244-135	Sequence 135, f
35	560.5	28.9	326	12	US-10-290-386-75	Sequence 75, f
36	560.5	28.9	326	12	US-10-290-386-135	Sequence 135, f
37	560.5	28.9	326	14	US-10-033-297-75	Sequence 75, f
38	560.5	28.9	326	14	US-10-033-297-135	Sequence 135, f
39	560	28.9	356	11	US-09-940-244-389	Sequence 389, f
40	542	28.0	335	11	US-09-940-244-149	Sequence 149, f
41	541.5	27.9	881	11	US-09-864-036A-560	Sequence 560, f
42	541.5	27.9	881	11	US-09-864-036A-566	Sequence 566, f
43	541.5	27.9	881	11	US-09-864-036A-568	Sequence 568, f
44	541.5	27.9	881	12	US-10-084-039-560	Sequence 560, f
45	541.5	27.9	881	12	US-10-084-039-566	Sequence 566, f

## ALIGNMENTS

RESULT 1	US-09-805-311-2	Application US/09805311
1	Sequence 2, Appl1	/cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2	Patent No. US20010034886A1	
3	GENERAL INFORMATION:	
4	APPLICANT: Mahajan, Pramod B.	
5	TITLE OF INVENTION: Rad2/FEN-1 Orthologues and Uses	
6	TITLE OF INVENTION: Theoret	
7	FILE REFERENCE: 09610	
8	CURRENT APPLICATION NUMBER: US/09/905,311	
9	CURRENT FILING DATE: 2001-03-13	
10	PRIOR APPLICATION NUMBER: 09/426,557	
11	PRIOR FILING DATE: 1999-10-22	
12	PRIOR APPLICATION NUMBER: 60/212,342	
13	PRIOR FILING DATE: 1998-12-15	
14	NUMBER OF SEQ ID NOS: 10	
15	SOFTWARE: FASTSP for Windows Version 1.0	
16	SEQ ID NO: 2	
17	LENGTH: 379	
18	TYPE: PRT	
19	ORGANISM: Zea mays	
20	US-09-805-311-2	
21	Query Match	100.0% Score 1939, DB 9, Length 379
22	Best Local Similarity	100.0% Pct. No. 1, 7e-174
23	Matches 379: Conservative	0: Mismatches 0: Indels 0: Gaps
24	1	1
25	1	1
26	1	1
27	1	1
28	1	1
29	1	1
30	1	1
31	1	1
32	1	1
33	1	1
34	1	1
35	1	1
36	1	1
37	1	1
38	1	1
39	1	1
40	1	1
41	1	1
42	1	1
43	1	1
44	1	1
45	1	1



```

QY 181 MSLSFGAPRLRLMDPSSKRIPIWEDVAVKVELELITMDQFIDLCILGCGYCSIK 240
DB 181 MSLSFGAPRLRLMDPSSKRIPIWEDVAVKVELELITMDQFIDLCILGCGYCSIK 240
QY 241 GIGGOTALKLIRHOGSISILENLKDKRYQIPEDMPYOEARLKEPNVTLIDPELKWTA 300
DB 241 GIGGOTALKLIRHOGSISILENLKDKRYQIPEDMPYOEARLKEPNVTLIDPELKWTA 300
QY 301 PDEGLISFLVKGNGFNEDRYTKAIEKISAKNKSOGRLSEFFKPTATTSAPLKRKETS 360
DB 301 PDEGLISFLVKGNGFNEDRYTKAIEKISAKNKSOGRLSEFFKPTATTSAPLKRKETS 360
QY 361 DKTSKAANKKTKAGGKKK 379
DB 361 DKTSKAANKKTKAGGKKK 379

```

## RESULT 2

```

US-09-805-311-6
Sequence 6, Application US/09805311
Patent No. US20010034886A1
GENERAL INFORMATION:
APPLICANT: Mahajan, Pramod B.
TITLE OF INVENTION: Rad2/FEN-1 Orthologues and Uses
FILE REFERENCE: 0961D
CURRENT APPLICATION NUMBER: US/09/805,311
PRIOR FILING DATE: 2001-03-13
PRIOR APPLICATION NUMBER: 09/426,557
PRIOR FILING DATE: 1999-10-22
PRIOR APPLICATION NUMBER: 60/112,332
PRIOR FILING DATE: 1998-12-15
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 6
LENGTH: 379
TYPE: PRT
ORGANISM: Zea mays
US-09-805-311-6

```

```

Query Match          100.0%: Score 1939; DB 9; Length 379;
Best Local Similarity 100.0%: Pred. No. 1,70-174;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MGKGLTKLADNAPRAKKEOKFESEYFGQRKIAVDASMSIYQFLVVGRTGNETLTENAGE 60
DB 1 MGKGLTKLADNAPRAKKEOKFESEYFGQRKIAVDASMSIYQFLVVGRTGNETLTENAGE 60
QY 61 VTSHLQGMENPTIRLEAGIKPVYVFDGKPPDMKKQELAKRYSKRDPAITKDLTEAVEYGD 120
DB 61 VTSHLQGMENPTIRLEAGIKPVYVFDGKPPDMKKQELAKRYSKRDPAITKDLTEAVEYGD 120
QY 121 KAIEKLSKRTYKVRQHNEDCKRLRLKMGVYVAVPSEAEACALCINXKVAVASSED 180
DB 121 KAIEKLSKRTYKVRQHNEDCKRLRLKMGVYVAVPSEAEACALCINXKVAVASSED 180
QY 181 MSLSFGAPRLRLMDPSSKRIPIWEDVAVKVELELITMDQFIDLCILGCGYCSIK 240
DB 181 MSLSFGAPRLRLMDPSSKRIPIWEDVAVKVELELITMDQFIDLCILGCGYCSIK 240
QY 241 GIGGOTALKLIRHOGSISILENLKDKRYQIPEDMPYOEARLKEPNVTLIDPELKWTA 300
DB 241 GIGGOTALKLIRHOGSISILENLKDKRYQIPEDMPYOEARLKEPNVTLIDPELKWTA 300
QY 301 PDEGLISFLVKGNGFNEDRYTKAIEKISAKNKSOGRLSEFFKPTATTSAPLKRKETS 360
DB 301 PDEGLISFLVKGNGFNEDRYTKAIEKISAKNKSOGRLSEFFKPTATTSAPLKRKETS 360
QY 361 DKTSKAANKKTKAGGKKK 379
DB 361 DKTSKAANKKTKAGGKKK 379

```

## RESULT 3

```

US-09-805-311-4
Sequence 4, Application US/09805311
Patent No. US20010034886A1
GENERAL INFORMATION:
APPLICANT: Mahajan, Pramod B.
TITLE OF INVENTION: Rad2/FEN-1 Orthologues and Uses
FILE REFERENCE: 0961D
CURRENT APPLICATION NUMBER: US/09/805,311
PRIOR FILING DATE: 2001-03-13
PRIOR FILING DATE: 1999-10-22
PRIOR APPLICATION NUMBER: 60/112,332
PRIOR FILING DATE: 1998-12-15
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 4
LENGTH: 379
TYPE: PRT
ORGANISM: Zea mays
US-09-805-311-4

```

```

Query Match          99.7%: Score 1933; DB 9; Length 379;
Best Local Similarity 99.7%: Pred. No. 6,20-174;
Matches 378; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 MGKGLTKLADNAPRAKKEOKFESEYFGQRKIAVDASMSIYQFLVVGRTGNETLTENAGE 60
DB 1 MGKGLTKLADNAPRAKKEOKFESEYFGQRKIAVDASMSIYQFLVVGRTGNETLTENAGE 60
QY 61 VTSHLQGMENPTIRLEAGIKPVYVFDGKPPDMKKQELAKRYSKRDPAITKDLTEAVEYGD 120
DB 61 VTSHLQGMENPTIRLEAGIKPVYVFDGKPPDMKKQELAKRYSKRDPAITKDLTEAVEYGD 120
QY 121 KAIEKLSKRTYKVRQHNEDCKRLRLKMGVYVAVPSEAEACALCINXKVAVASSED 180
DB 121 KAIEKLSKRTYKVRQHNEDCKRLRLKMGVYVAVPSEAEACALCINXKVAVASSED 180
QY 181 MSLSFGAPRLRLMDPSSKRIPIWEDVAVKVELELITMDQFIDLCILGCGYCSIK 240
DB 181 MSLSFGAPRLRLMDPSSKRIPIWEDVAVKVELELITMDQFIDLCILGCGYCSIK 240
QY 241 GIGGOTALKLIRHOGSISILENLKDKRYQIPEDMPYOEARLKEPNVTLIDPELKWTA 300
DB 241 GIGGOTALKLIRHOGSISILENLKDKRYQIPEDMPYOEARLKEPNVTLIDPELKWTA 300
QY 301 PDEGLISFLVKGNGFNEDRYTKAIEKISAKNKSOGRLSEFFKPTATTSAPLKRKETS 360
DB 301 PDEGLISFLVKGNGFNEDRYTKAIEKISAKNKSOGRLSEFFKPTATTSAPLKRKETS 360
QY 361 DKTSKAANKKTKAGGKKK 379
DB 361 DKTSKAANKKTKAGGKKK 379

```

## RESULT 4

```

US-09-805-311-8
Sequence 8, Application US/09805311
Patent No. US20010034886A1
GENERAL INFORMATION:
APPLICANT: Mahajan, Pramod B.
TITLE OF INVENTION: Rad2/FEN-1 Orthologues and Uses
FILE REFERENCE: 0961D
CURRENT APPLICATION NUMBER: US/09/805,311
PRIOR FILING DATE: 2001-03-13
PRIOR FILING DATE: 1999-10-22
PRIOR APPLICATION NUMBER: 60/112,332
PRIOR FILING DATE: 1998-12-15
NUMBER OF SEQ ID NOS: 10

```

SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 8  
LENGTH: 379  
TYPE: PRT  
ORGANISM: Zea mays  
US-09-805-311-8

Query Match  
Best Local Similarity 99.7%; P-adj. No. 6.2e-174;  
Matches 378; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGKGLTKLADNAPKAKKEQFESEYGRKIAVDASMSIYQFLIVSGTGMETLTNEAG 60  
DB 1 MGKGLTKLADNAPKAKKEQFESEYGRKIAVDASMSIYQFLIVSGTGMETLTNEAG 60  
QY 61 VTSHLQGMFNRTIRLEAGIKPVYVFDGKPPDMKKQELAKRSKRDATKDLTAVVEGD 120  
DB 61 VTSHLQGMFNRTIRLEAGIKPVYVFDGKPPDMKKQELAKRSKRDATKDLTAVVEGD 120  
QY 121 KDAIEKLSKRTYKVTROHNEDECKRLRLMGVYVAFSPSEAEACALCTNKREAVASGD 180  
DB 121 KDAIEKLSKRTYKVTROHNEDECKRLRLMGVYVAFSPSEAEACALCTNKREAVASGD 180  
QY 181 KDSLTFFGAPRFLRLHLDSPSSKRIPIVMEFDVAKVLEELTLTQDFDLCTLGGDVCESIK 240  
DB 181 KDSLTFFGAPRFLRLHLDSPSSKRIPIVMEFDVAKVLEELTLTQDFDLCTLGGDVCESIK 240  
QY 241 GIGGOTALKLIRQHGSIESTLENLKKRYQIPEMPPYQEARLREPVATLDPELKMTA 300  
DB 241 GIGGOTALKLIRQHGSIESTLENLKKRYQIPEMPPYQEARLREPVATLDPELKMTA 300  
QY 301 PDEGLISFLVKNQFNEDRYTAKLEIKSAKNSQGRLESFETATTSAPLRKETS 360  
DB 301 PDEGLISFLVKNQFNEDRYTAKLEIKSAKNSQGRLESFETATTSAPLRKETS 360  
QY 361 DKTSAANKKTAKGKKK 379  
DB 361 DKTSAANKKTAKGKKK 379

## RESULT 5

US-09-805-311-10  
Sequence 10, Application US/09805311  
Patent No. US20010034886A1  
GENERAL INFORMATION:  
APPLICANT: Mahajan, Pramod B.  
TITLE OF INVENTION: Rad2/FEN-1 Orthologues and Uses  
FILE REFERENCE: 0961D  
CURRENT APPLICATION NUMBER: US/09/805,311  
PRIOR FILING DATE: 2001-03-13  
PRIOR APPLICATION NUMBER: 09/426,557  
PRIOR FILING DATE: 1999-10-22  
PRIOR APPLICATION NUMBER: 60/112,332  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 10  
LENGTH: 380  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-805-311-10

Query Match  
Best Local Similarity 53.2%; Score 1032; DB 9; Length 380;  
Matches 206; Conservative 69; Mismatches 98; Indels 10; Gaps 6;

QY 1 MGKGLTKLADNAPKAKKEQFESEYGRKIAVDASMSIYQFLIVSGTGMETLTNEAG 60  
DB 1 MGKGLTKLADNAPKAKKEQFESEYGRKIAVDASMSIYQFLIVSGTGMETLTNEAG 60  
QY 61 VTSHLQGMFNRTIRLEAGIKPVYVFDGKPPDMKKQELAKRSKRDATKDLTAVVEGD 120  
DB 61 VTSHLQGMFNRTIRLEAGIKPVYVFDGKPPDMKKQELAKRSKRDATKDLTAVVEGD 120

DB 60 TISHLMKQEVRTIRKMNCKIKPVYVFDGKPPDQLKSGHLAKRSERAEAKKQJQQAQAGA  
QY 121 KDAIEKLSKRTYKVTROHNEDECKRLRLMGVYVAFSPSEAEACALCTNKREAVASGD 180  
DB 121 KDAIEKLSKRTYKVTROHNEDECKRLRLMGVYVAFSPSEAEACALCTNKREAVASGD 180  
QY 181 KDSLTFFGAPRFLRLHLDSPSSKRIPIVMEFDVAKVLEELTLTQDFDLCTLGGDVCESIK 240  
DB 181 KDSLTFFGAPRFLRLHLDSPSSKRIPIVMEFDVAKVLEELTLTQDFDLCTLGGDVCESIK 240  
QY 241 GIGGOTALKLIRQHGSIESTLENLKKRYQIPEMPPYQEARLREPVATLDPELKMTA 300  
DB 241 GIGGOTALKLIRQHGSIESTLENLKKRYQIPEMPPYQEARLREPVATLDPELKMTA 300  
QY 301 PDEGLISFLVKNQFNEDRYTAKLEIKSAKNSQGRLESFETATTSAPLRKETS 360  
DB 301 PDEGLISFLVKNQFNEDRYTAKLEIKSAKNSQGRLESFETATTSAPLRKETS 360  
QY 361 DKTSAANKKTAKGKKK 379  
DB 361 DKTSAANKKTAKGKKK 379

## RESULT 6

US-09-940-244-137  
Sequence 137, Application US/09940244  
Publication No. US20040044796A1  
GENERAL INFORMATION:  
APPLICANT: Netti, Bruce P.  
APPLICANT: Hall, Jeff G.  
APPLICANT: Lyamichay, Victor  
APPLICANT: Smith, Lloyd M.  
TITLE OF INVENTION: Reactions on Dendrimers  
FILE REFERENCE: PORS-06476  
CURRENT APPLICATION NUMBER: US/09/940,244  
PRIOR FILING DATE: 2002-05-06  
NUMBER OF SEQ ID NOS: 422  
SOFTWARE: patentin version 3.1  
SEQ ID NO 137  
LENGTH: 380  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-940-244-137

Query Match  
Best Local Similarity 53.8%; P-adj. No. 7.2e-89;  
Matches 206; Conservative 69; Mismatches 98; Indels 10; Gaps 6;

QY 1 MGKGLTKLADNAPKAKKEQFESEYGRKIAVDASMSIYQFLIVSGTGMETLTNEAG 60  
DB 1 MGKGLTKLADNAPKAKKEQFESEYGRKIAVDASMSIYQFLIVSGTGMETLTNEAG 60  
QY 61 VTSHLQGMFNRTIRLEAGIKPVYVFDGKPPDMKKQELAKRSKRDATKDLTAVVEGD 120  
DB 61 VTSHLQGMFNRTIRLEAGIKPVYVFDGKPPDMKKQELAKRSKRDATKDLTAVVEGD 120  
QY 121 KDAIEKLSKRTYKVTROHNEDECKRLRLMGVYVAFSPSEAEACALCTNKREAVASGD 180  
DB 121 KDAIEKLSKRTYKVTROHNEDECKRLRLMGVYVAFSPSEAEACALCTNKREAVASGD 180  
QY 181 KDSLTFFGAPRFLRLHLDSPSSKRIPIVMEFDVAKVLEELTLTQDFDLCTLGGDVCESIK 240  
DB 181 KDSLTFFGAPRFLRLHLDSPSSKRIPIVMEFDVAKVLEELTLTQDFDLCTLGGDVCESIK 240  
QY 241 GIGGOTALKLIRQHGSIESTLENLKKRYQIPEMPPYQEARLREPVATLDPELKMTA 300  
DB 241 GIGGOTALKLIRQHGSIESTLENLKKRYQIPEMPPYQEARLREPVATLDPELKMTA 300  
QY 301 PDEGLISFLVKNQFNEDRYTAKLEIKSAKNSQGRLESFETATTSAPLRKETS 360  
DB 301 PDEGLISFLVKNQFNEDRYTAKLEIKSAKNSQGRLESFETATTSAPLRKETS 360  
QY 361 DKTSAANKKTAKGKKK 379  
DB 361 DKTSAANKKTAKGKKK 379



```

QY 1 MGICG:IKLLADNAPKAMKEQKFESEYFGKRIAVDASMSIYOFLLIYVGR-3MPTL-7PAGE 60
DB 1 MGICG:IKLLADNAPKAMKEQKFESEYFGKRIAVDASMSIYOFLLIYVGR-3MPTL-7PAGE 59
QY 61 VTSILOGMENRTIRLEAGIKPVYVFDGKPPDMKKOELAKRYSKROCATKOLTEAVEVGD 120
DB 60 TSHLMGMFYRTIRWMENGIKPVYVFDGKPPDMKKOELAKRYSKROCATKOLTEAVEVGD 119
QY 121 KDAIEKLSKRYKVRQIHNEDECKRLRLNGVAVVAVAPSAEACALCINDKVFVAASED 180
DB 120 EQVEVEKTRIAVKVRQIHNEDECKRLRLNGVAVVAVAPSAEACALCINDKVFVAASED 179
QY 181 MDLTFGAPRFLRHIMDPSSKKIPVMEFVAVKLEBELTMDQFCTLITGGCOWGSSIK 240
DB 180 MDLTFGAPRFLRHIMDPSSKKIPVMEFVAVKLEBELTMDQFCTLITGGCOWGSSIK 239
QY 241 GIGGOTALKLIRQHSIESILENLKRDYQIPEDPYQAPRLSEKPNVYTD-1PECKK 296
DB 240 GIGGOTALKLIRQHSIESILENLKRDYQIPEDPYQAPRLSEKPNVYTD-1PECKK 298
QY 299 TAPDEGLISFLVCKGFMEDRVTAKIEKIKSAKKSSGGLSEPFKRTATISAPLKKE 358
DB 298 TAPDEGLISFLVCKGFMEDRVTAKIEKIKSAKKSSGGLSEPFKRTATISAPLKKE 357
QY 359 TSDK--TSKAANKKTKAGCKK 379
DB 358 TSDK--TSKAANKKTKAGCKK 377

```

```

RESULT 9
US-09-940-244-138
Sequence 138, Application US/09940244
Publication No. US20030044796A1
GENERAL INFORMATION:
APPLICANT: Neti, Bruce P.
APPLICANT: Hall, Jeff G.
APPLICANT: Lyamichev, Victor
APPLICANT: Smith, Lloyd M.
TITLE OF INVENTION: Reactions on dendrimers
FILE REFERENCE: FORS-06478
CURRENT APPLICATION NUMBER: US/09/940,244
CURRENT FILING DATE: 2002-05-06
NUMBER OF SEQ ID NOS: 422
SOFTWARE: PatentIn version 3.1
LENGTH: 378
TYPE: PRI
ORGANISM: Mus musculus
US-09-940-244-138

```

```

Query Match 52.3%; Score 1014; DB 11; Length 378;
Best Local Similarity 53.2%; Pred. No. 3,6e-87;
Matches 202; Conservative 59; Mismatches 101; Indels 8; Gaps 6;

```

```

QY 1 MGICG:IKLLADNAPKAMKEQKFESEYFGKRIAVDASMSIYOFLLIYVGR-3MPTL-7PAGE 60
DB 1 MGICG:IKLLADNAPKAMKEQKFESEYFGKRIAVDASMSIYOFLLIYVGR-3MPTL-7PAGE 59
QY 61 VTSILOGMENRTIRLEAGIKPVYVFDGKPPDMKKOELAKRYSKROCATKOLTEAVEVGD 120
DB 60 TSHLMGMFYRTIRWMENGIKPVYVFDGKPPDMKKOELAKRYSKROCATKOLTEAVEVGD 119
QY 121 KDAIEKLSKRYKVRQIHNEDECKRLRLNGVAVVAVAPSAEACALCINDKVFVAASED 180
DB 120 EQVEVEKTRIAVKVRQIHNEDECKRLRLNGVAVVAVAPSAEACALCINDKVFVAASED 179
QY 181 MDLTFGAPRFLRHIMDPSSKKIPVMEFVAVKLEBELTMDQFCTLITGGCOWGSSIK 240
DB 180 MDLTFGAPRFLRHIMDPSSKKIPVMEFVAVKLEBELTMDQFCTLITGGCOWGSSIK 239
QY 241 GIGGOTALKLIRQHSIESILENLKRDYQIPEDPYQAPRLSEKPNVYTD-1PECKK 296
DB 240 GIGGOTALKLIRQHSIESILENLKRDYQIPEDPYQAPRLSEKPNVYTD-1PECKK 298
QY 299 TAPDEGLISFLVCKGFMEDRVTAKIEKIKSAKKSSGGLSEPFKRTATISAPLKKE 358
DB 298 TAPDEGLISFLVCKGFMEDRVTAKIEKIKSAKKSSGGLSEPFKRTATISAPLKKE 357

```

```

QY 300 APDEEGISFLVCKGFMEDRVTAKIEKIKSAKKSSGGLSEPFKRTATISAPLKKE 358
DB 298 APDEEGISFLVCKGFMEDRVTAKIEKIKSAKKSSGGLSEPFKRTATISAPLKKE 357
QY 357 TSDK--TSKAANKKTKAGCKK 379
DB 356 TSDK--TSKAANKKTKAGCKK 378

```

```

RESULT 10
US-10-290-386-138
Sequence 138, Application US/10290386
Publication No. US2003015297A1
GENERAL INFORMATION:
APPLICANT: Neti, Bruce P.
APPLICANT: Hall, Jeff G.
APPLICANT: Lyamichev, Victor
APPLICANT: Smith, Lloyd M.
TITLE OF INVENTION: Reactions on dendrimers
FILE REFERENCE: FORS-07459
CURRENT APPLICATION NUMBER: US/10/290,386
CURRENT FILING DATE: 2002-11-07
PRIOR APPLICATION NUMBER: 09/7713,602
PRIOR FILING DATE: 2000-11-15
PRIOR APPLICATION NUMBER: 09/381,212
PRIOR FILING DATE: 2000-02-08
PRIOR APPLICATION NUMBER: 09/350,335
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 08/323,516
PRIOR FILING DATE: 1997-03-24
PRIOR APPLICATION NUMBER: 08/759,018
PRIOR FILING DATE: 1996-12-02
PRIOR APPLICATION NUMBER: 08/756,386
PRIOR FILING DATE: 1996-11-26
PRIOR APPLICATION NUMBER: 08/662,853
PRIOR FILING DATE: 1996-07-12
PRIOR APPLICATION NUMBER: 08/599,491
PRIOR FILING DATE: 1996-01-24
NUMBER OF SEQ ID NOS: 253
SOFTWARE: PatentIn version 3.1
LENGTH: 378
TYPE: PRI
ORGANISM: Artificial
FEATURE:
OTHER INFORMATION: Synthetic
US-10-290-386-138

```

```

Query Match 52.3%; Score 1014; DB 12; Length 378;
Best Local Similarity 53.2%; Pred. No. 3,6e-87;
Matches 202; Conservative 59; Mismatches 101; Indels 8; Gaps 6;

```

```

QY 1 MGICG:IKLLADNAPKAMKEQKFESEYFGKRIAVDASMSIYOFLLIYVGR-3MPTL-7PAGE 60
DB 1 MGICG:IKLLADNAPKAMKEQKFESEYFGKRIAVDASMSIYOFLLIYVGR-3MPTL-7PAGE 59
QY 61 VTSILOGMENRTIRLEAGIKPVYVFDGKPPDMKKOELAKRYSKROCATKOLTEAVEVGD 120
DB 60 TSHLMGMFYRTIRWMENGIKPVYVFDGKPPDMKKOELAKRYSKROCATKOLTEAVEVGD 119
QY 121 KDAIEKLSKRYKVRQIHNEDECKRLRLNGVAVVAVAPSAEACALCINDKVFVAASED 180
DB 120 EQVEVEKTRIAVKVRQIHNEDECKRLRLNGVAVVAVAPSAEACALCINDKVFVAASED 179
QY 181 MDLTFGAPRFLRHIMDPSSKKIPVMEFVAVKLEBELTMDQFCTLITGGCOWGSSIK 240
DB 180 MDLTFGAPRFLRHIMDPSSKKIPVMEFVAVKLEBELTMDQFCTLITGGCOWGSSIK 239
QY 241 GIGGOTALKLIRQHSIESILENLKRDYQIPEDPYQAPRLSEKPNVYTD-1PECKK 296
DB 240 GIGGOTALKLIRQHSIESILENLKRDYQIPEDPYQAPRLSEKPNVYTD-1PECKK 298
QY 299 TAPDEGLISFLVCKGFMEDRVTAKIEKIKSAKKSSGGLSEPFKRTATISAPLKKE 358
DB 298 TAPDEGLISFLVCKGFMEDRVTAKIEKIKSAKKSSGGLSEPFKRTATISAPLKKE 357

```



```

Db 178 MDLCTPTPLNHLTFSEAKKEPIHEIDTEIVLRGLDLEOFVDCIMLGDYCESTR 237
Qy 241 GIGGOTALKLIRHSHSIESLENNL-----NKDRYOIPEDMPYQEARLFEKPNVLDLPE 235
Db 238 GAGPVATLKLIRHSHSIESLENNL-----NKDRYOIPEDMPYQEARLFEKPNVLDLPE 235
Qy 296 --LKWTAPEDEGLISFLVKNQNGENEDRVTKALEIKSAKNSQGRLESFFKPTATT--- 240
Db 297 INLWSPKPEKELIYLCDDKKPSERFVKSISRLKGLKSGIOGRDGFQVVFVKTRQ 246
Qy 351 -SAPLKKRETSQKTSKAANKKTK 373
Db 357 LAAAKKRAOENKRLNK-NKKNVTK 379

RESULT 13
US-10-290-366-139
Sequence 139, Application US/10290386
Publication No. US20030152971A1
GENERAL INFORMATION:
APPLICANT: Lyamichiev, Victor
APPLICANT: Hall, Jeff G.
APPLICANT: Lukowiak, Andrew A.
TITLE OF INVENTION: Methods and Compositions for Detecting Target Sequences
FILE REFERENCE: FORS-07459
CURRENT APPLICATION NUMBER: US/10/290,386
PRIOR APPLICATION NUMBER: 2002-11-07
PRIOR FILING DATE: 2002-02-27
PRIOR APPLICATION NUMBER: 60/344,946
PRIOR FILING DATE: 2001-11-07
PRIOR APPLICATION NUMBER: 09/713,601
PRIOR FILING DATE: 2000-11-15
PRIOR APPLICATION NUMBER: 08/381,212
PRIOR FILING DATE: 2000-02-08
PRIOR APPLICATION NUMBER: 09/350,309
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 08/823,516
PRIOR FILING DATE: 1997-03-24
PRIOR APPLICATION NUMBER: 08/759,038
PRIOR FILING DATE: 1996-12-02
PRIOR APPLICATION NUMBER: 08/756,386
PRIOR FILING DATE: 1996-11-26
PRIOR APPLICATION NUMBER: 08/682,853
PRIOR FILING DATE: 1996-07-12
PRIOR APPLICATION NUMBER: 08/559,491
PRIOR FILING DATE: 1996-01-24
NUMBER OF SEQ ID NOS: 253
SOFTWARE: Patent version 3.1
SEQ ID NO 139
LENGTH: 362
TYPE: PKT
ORGANISM: Artificial
FEATURE:
OTHER INFORMATION: Synthetic
US-10-290-386-139

Query Match 48.4%; Score 938; DS 12; Length 362;
Best Local Similarity 49.5%; Pred. No. 5,5e-60;
Matches 190; Conservative 69; Mismatches 109; Indels 16; Gaps 6;

```

```

Db 121 K---MKOERLVKSENEENEOKULJLGMGIPTTAPTEAEOCAELAKKGVYAAASED 1
Qy 181 MDLCTGAPRLRHH-MUPSSKKIPWKEFVAVVLELELTMQOFTDCLTGGCYCDSIK 2
Db 178 MDLCTGAPRLRHH-MUPSSKKIPWKEFVAVVLELELTMQOFTDCLTGGCYCDSIK 2
Qy 241 GIGGOTALKLIRHSHSIESLENNL-----NKDRYOIPEDMPYQEARLFEKPNVLDLPE 2
Db 238 GAGPVATLKLIRHSHSIESLENNL-----NKDRYOIPEDMPYQEARLFEKPNVLDLPE 2
Qy 296 --LKWTAPEDEGLISFLVKNQNGENEDRVTKALEIKSAKNSQGRLESFFKPTATT--- 2
Db 297 INLWSPKPEKELIYLCDDKKPSERFVKSISRLKGLKSGIOGRDGFQVVFVKTRQ 2
Qy 351 -SAPLKKRETSQKTSKAANKKTK 373
Db 357 LAAAKKRAOENKRLNK-NKKNVTK 379

RESULT 14
US-10-033-297-139
Sequence 139, Application US/10033297
Publication No. US20020187486A1
GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
APPLICANT: Lyamichiev, Victor I.
APPLICANT: Brown, Mary Ann B.
TITLE OF INVENTION: Detection of Nucleic Acids By Multiple
Sequential Invasive Cleavages
NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESS:
ADDRESS: Medien & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC/OS/MS-DOS
SOFTWARE: Patent In Re: Case #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/033,297
FILING DATE: 12-NOV-02
CLASSIFICATION: <UNKNOWN>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/350,597
FILING DATE: 09-JUL-1999
APPLICATION NUMBER: US/04/823,516
FILING DATE: 24-MAR-1997
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 04/758,314
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/756,186
FILING DATE: 29-NOV-1996
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/559,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Lyamichiev, Victor I.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 139:
SEQUENCE CHARACTERISTICS:

```

LENGTH: 382 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: No. US20020187486A1 Relevant  
 TOPOLOGY: No. US20020187486A1 Relevant  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 139  
 US-10-033-297-139

Query Match 48.4% Score 938; DB 14; Length 382;  
 Best Local Similarity 49.5%; Pred. No. 5,56-80;  
 Matches 190; Conservative 69; Mismatches 105; Indels 16; Gaps 6;

QY 1 MGICGLTLADNAKAKKFKFESYFGRKLAVDASXSIYQFLVGRGTGNETLTKRGE 60  
 DB 1 MGICGLNLIISFHVSAIRKSDIKSFGRKVAIDASMSLYQFLVAVRQDQGLTKRGE 60  
 QY 61 VTSEJGQFNRITRLLENGIKRVYVFDGKPRDMKQGLAKRYSKDDTKDLTEAVTGD 120  
 DB 61 TTSRLKMFYRLKMDGIRKQYVBDKPPDLKSHLTKRSSRVESEKRIKVAATTIRK 120  
 QY 121 KDAIEKSKRTYKVRQHNEDSKRLRLMGVYVPAEAEAFCAALCINKEVAVASED 180  
 DB 121 K--KQERRLVKSKHNEEOKLGLKGIPIYIAPTEAAGCALAKKRVAAASRD 177  
 QY 181 MSLITGAPRFLRLHMDSSKKIRPMEDVAKVLEDELIMDQFLDLCIGGQDYCSIK 240  
 DB 178 MDLTGRTPLRLHLESEAKKEPHEIDTELVLKGLITTECFVDLTKLQDYCESIR 237  
 QY 241 GIGGQATLKLIRHGSISILENC-----SKDRQIPEDNPQEARL-FKEPNVLDIPE 285  
 DB 238 GVGPTALKLIKTRHSTIKETVPIESGESNNTKAIPEQWPKQAKMLJLDPNVLDENE 286  
 QY 296 --LKWTADEGLISFLVKGNGFNEDRYKALEIKTSAKNKSQGRLESEFFKPA--ATT-- 350  
 DB 297 TLKWSPPREKELTEYLDODKKSESEKSGISRLKGLKSGIQGRLDGFFGVVFKTKEQ 356  
 QY 351 -SAPLRKRTSDKTSKAANKTK 373  
 DB 357 IAAAKRAQENKLNK-NKNVTK 379

## RESULT 15

US-09-777-430A-27  
 Sequence 27, Application US/C9777430A  
 Patent No. US20020128465A1  
 GENERAL INFORMATION:  
 APPLICANT: Lyamichev, Victor  
 APPLICANT: Skrzypczynski, Zbigniew  
 APPLICANT: Allawi, Hatim T.  
 APPLICANT: Mayland, Sarah R.  
 APPLICANT: Yakova, Tsetska  
 TITLE OF INVENTION: Charge Tags and the Separation of Nucleic Acid Molecules  
 FILE REFERENCE: FORS-04912  
 CURRENT APPLICATION NUMBER: US/09/777,430A  
 CURRENT FILING DATE: 2001-02-06  
 NUMBER OF SEQ ID NOS: 85  
 SOFTWARE: Patentin version 3.1  
 SEQ ID NO 27  
 LENGTH: 340  
 TYPE: PRT  
 ORGANISM: Pyrococcus furiosus  
 US-09-777-430A-27

Query Match 33.4% Score 648.5; DB 10; Length 340;  
 Best Local Similarity 41.3%; Pred. No. 9,96-53;  
 Matches 142; Conservative 62; Mismatches 107; Indels 33; Gaps 0;

QY 19 KFOKESTFGRKLAVDASXSIYQFLVGRGTGNETLTKRGEVYVPAEAEAFCAALCINKEVAVASED 34  
 DB 12 KELEENLTKGKIALVDALNATVQFLSTIRQKIDGTPCLDKSKGRLTSHLSGLFYRIININKEA 71  
 QY 79 GIKPYVVFDEGPEPKKKLEKRRREAREEAEENKREALENGEILEAKRYAQARTRVNEML 138

DB 72 GIKPYVVFDEGPEPKKKLEKRRREAREEAEENKREALENGEILEAKRYAQARTRVNEML 1  
 QY 139 NEDCKPLRLMGVYVPAEAEAFCAALCINKEVAVASEDMDSLTFGAPRFLRHMDP 1  
 DB 132 TFDKKLLELMGIPYVAPSEGEAOAAVMAKGSYVASASQDUSLTFGAPRFLRHMDTT 1  
 QY 199 SSKRP-----VSEFDVAKVLEDELIMDQFLDLCIGGQDY-CDSITKGTG 2  
 DB 192 GKRLTPGKNYVVEIKRPLILR-----EVLKELKLTFSKLELAIVGTDVNPQIKIGL 2  
 QY 245 CTALKLIRHGSISILENLRKD--RYQIPETMPYQEARMLKEPNVTLDIPLKATAP 3  
 DB 248 KKALEIVRH-----SKDPLAKFGKQSDVCLYAIKKEFFLNPVT-FNRYNIVMRD 2  
 QY 302 LEEELISFLVKGNGFNEDRYKALEIKTSAKNKSQGRLESEFFK 345  
 DB 296 DEKDLKFLDDEHDSSEKRVKSGERLKLKAIKSKUSTIDSMER 359

Search completed: October 22, 2003, 12:44:18  
 Job time : 32 secs